IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

VOCALIFE LLC,	)(	
PLAINTIFF,	)(	CIVIL ACTION NO.
	)(	2:19-CV-123-JRG
VS.	)(	MARSHALL, TEXAS
	)(	
AMAZON.COM, INC. and	)(	
AMAZON.COM LLC,	)(	OCTOBER 5, 2020
DEFENDANTS.	)(	8:30 A.M.

TRANSCRIPT OF JURY TRIAL
MORNING SESSION
BEFORE THE HONORABLE JUDGE RODNEY GILSTRAP
UNITED STATES CHIEF DISTRICT JUDGE

FOR THE PLAINTIFF:

MR. ALFRED R. FABRICANT
MR. PETER LAMBRIANAKOS
MR. VINCENT J. RUBINO, III
MS. AMY PARK
MR. ENRIQUE ITURRALDE
FABRICANT LLP
230 Park Avenue, 3rd Floor W.
New York, NY 10169

MR. SAMUEL F. BAXTER
MS. JENNIFER L. TRUELOVE
MCKOOL SMITH, P.C.
104 East Houston Street, Suite 300
Marshall, TX 75670

FOR THE DEFENDANTS:

MR. JOSEPH R. RE
ALAN G. LAQUER
KENDALL M. LOEBBAKA
JOSHUA J. STOWELL
KNOBBE, MARTENS, OLSON & BEAR, LLP
2040 Main Street, Fourteenth Floor
Irvine, CA 92614

MR. COLIN B. HEIDEMAN
KNOBBE, MARTENS, OLSON & BEAR, LLP
925 Fourth Avenue, Suite 2500
Seattle, WA 98104

MS. JENNIFER H. DOAN
MR. JOSHUA R. THANE
MR. KYLE R. AKIN
HALTOM & DOAN
6500 Summerhill Road, Suite 100
Texarkana, TX 75503

MR. J. DAVID HADDEN
MR. RAVI RANGANATH
MR. THOMAS JOHN FOX
FENWICK & WEST LLP
801 California Street
Mountain View, CA 94041

MR. DERON R. DACUS
THE DACUS FIRM, PC
821 ESE Loop 323, Suite 430
Tyler, TX 75701

COURT REPORTER: Shelly Holmes, CSR, TCRR
Official Reporter
United States District Court
Eastern District of Texas
Marshall Division
100 E. Houston Street
Marshall, Texas 75670
(903) 923-7464

(Proceedings recorded by mechanical stenography, transcript produced on a CAT system.)

P R O C E E D I N G S

(Jury out.)

COURT SECURITY OFFICER: All rise.

THE COURT: Be seated, please.

Are the parties prepared to read into the record those items from the list of pre-admitted exhibits used during Friday's portion of the trial?

MR. AKIN: Yes, Your Honor.

MS. PARK: Yes, Your Honor.

THE COURT: Please proceed.

MS. PARK: Amy Park for Plaintiff, Vocalife.

The following exhibits were used on Friday and are being moved into evidence: PTX-109, PTX-110, PTX-111, and PTX-130.

THE COURT: Any objection to that rendition from the Defendants?

MR. AKIN: No objection.

THE COURT: Do Defendants have a similar rendition to read into the record?

MR. AKIN: Yes, Your Honor. Kyle Akin on behalf of Amazon.

Amazon moves the following exhibits into evidence that were used last Friday, October 2nd: DTX-1, DTX-14, DTX-685, DTX-689, DTX-980, and DTX-980A.

THE COURT: All right. Any objection to that

08:31:32 rendition by the Plaintiff?
08:31:35 MS. PARK: No objection.
08:31:36 THE COURT: Thank you, counsel.
08:31:41 We ended Friday with Mr. McAlexander on the witness stand.
08:31:47 Is Mr. McAlexander present?
08:31:50 If you would, please, return to the witness stand, sir.
08:31:53 And, Mr. Rubino, were we in the middle of your direct?
08:31:57 MR. RUBINO: Yes, Your Honor.
08:31:58 THE COURT: If you'd like to go to the podium and prepare to continue.
08:32:02 And while he's doing that, Mr. Johnston, if you'd bring in the jury.
08:32:48 COURT SECURITY OFFICER: All rise.
08:32:51 (Jury in.)
08:32:53 THE COURT: Welcome back, ladies and gentlemen of the jury. Please have a seat.
08:33:14 When we recessed on Friday evening, Mr. McAlexander was testifying as Plaintiff's technical expert witness. We'll continue with the direct examination of Mr. McAlexander by the Plaintiff.
08:33:33 Mr. Rubino, you may continue your direct examination.

08:33:35 JOSEPH MCALEXANDER, III, PLAINTIFF'S WITNESS,
08:33:35 PREVIOUSLY SWORN
08:33:35 DIRECT EXAMINATION CONTINUED
08:33:38 BY MR. RUBINO:
08:33:38 Q. Good morning, Mr. McAlexander.
08:33:40 A. Good morning, sir.
08:33:42 MR. RUBINO: If we could please, Mr. Thompson,
08:33:48 have Slide Demonstrative 10. Slide 10, please.
08:33:53 Q. (By Mr. Rubino) Mr. McAlexander, can you refresh us as
08:33:57 to what we were discussing on Friday regarding your
08:33:59 analysis?
08:34:00 A. Yes. We were just starting the claim construction
08:34:02 analysis, looking at Claim 1 of the '049 patent. And we
08:34:07 were generally describing what -- what is before you in
08:34:12 this slide is a demonstrative that recites the language
08:34:17 from the claim.
08:34:18 And I have specifically inserted the bracketed
08:34:21 letters [A] through [F] so that as I navigate step-by-step
08:34:26 through this claim, it would help the jury hopefully to
08:34:31 orient or can stay oriented as to exactly where we are in
08:34:37 the claim analysis.
08:34:38 MR. RUBINO: Mr. Thompson, if we could have
08:34:41 Slide 15, please?
08:34:42 Q. (By Mr. Rubino) And, Mr. McAlexander, if you could
08:34:45 refresh us as to what you were talking about on Friday with

08:34:49 regard to this slide?
08:34:50 A. Yes. With regard to Slide 15, I have identified the
08:34:52 listing of the accused Amazon products and listed in the
08:34:53 first row across the top, and these are all Echo-type
08:34:58 devices.
08:34:59 I've also indicated in the second row the
08:35:01 different generations that -- that are respectively
08:35:08 associated with each one of these products.
08:35:11 I have also in the lower row, the last row,
08:35:13 identified some pictures to give you some understanding of
08:35:15 what the look is -- look and feel of the product is.
08:35:19 And then in the center column, I've identified the
08:35:25 type of mic array -- microphone array that is found within
08:35:29 each one of these. And if we look across the products,
08:35:32 they go from four microphones to seven microphones to eight
08:35:37 microphones in some type of an array.
08:35:41 Q. And, Mr. McAlexander, did you have some physical units
08:35:44 that you had inspected, as well?
08:35:47 A. Yes.
08:35:52 MR. RUBINO: Your Honor, if I may approach the
08:35:54 witness with the physical exhibits.
08:35:55 THE COURT: You may approach the witness.
08:36:12 Q. (By Mr. Rubino) Mr. McAlexander, what are the numbers
08:36:14 of the exhibits that you're looking at currently?
08:36:16 A. The first is Plaintiff's Exhibit 654, which is a box

08:36:21 that includes the Echo device.
08:36:24 And the second is PTX-647, which is a teardown of
08:36:38 the Echo Dot Generation 3.
08:36:41 Q. Thank you, sir.
08:36:43 You can put those aside for the moment.
08:36:53 MR. RUBINO: May we please have PTX-15 -- 115?
08:36:59 THE COURT: Mr. Rubino, pull the microphone a
08:37:02 little closer to you, please. Thank you.
08:37:05 Q. (By Mr. Rubino) Sir, did you review any technical
08:37:17 documents in preparation or in rendering your opinions?
08:37:19 A. Yes, I did look at a number of technical documents that
08:37:22 were produced in this matter.
08:37:24 Q. Is this one of them?
08:37:25 A. PTX-115 is one of them, yes.
08:37:27 Q. And what type of document is this?
08:37:30 A. This document, if you'll notice at the bottom, is a
08:37:36 document that comes from a specific website, and so it
08:37:38 would be a document that's available to the public through
08:37:40 the normal website browser.
08:37:44 MR. RUBINO: If we could take a look at Page --
08:37:56 the second page of the document, please.
08:37:58 Q. (By Mr. Rubino) Can you explain what the top portion
08:38:01 of this document refers to?
08:38:02 A. Well, PTX-115, which we are looking at, is a document
08:38:13 that is entitled The Audio Front End Architecture For

08:38:20 Sonar. One of the types of product. And this is referred
08:38:23 to as the audio front end; the AFE is a name that we give
08:38:28 it sometimes.

08:38:29 And this represents a -- a block diagram that
08:38:33 shows the functional groupings of different aspects of
08:38:40 the -- the audio front end.

08:38:40 And, for instance, you'll notice with this
08:38:42 particular one, the top left corner, is -- is the
08:38:46 microphone array. And, in this case, it represents those
08:38:49 as dots. Each dot represents a mic.

08:38:55 And there are seven of them that's highlighted by
08:38:58 the number in -- in the arrow, that is just to the left --
08:39:01 or just to the right of the microphones.

08:39:04 And these microphones receive a signal. They
08:39:10 input it into a calibration block. It goes through a high
08:39:17 pass filter block. The high pass filter block is the one
08:39:20 that sets a ground floor of 80 hertz -- 80 hertz. And
08:39:20 anything that's below 80 hertz, it basically filters out.
08:39:25 It allows bypass -- allows those signals above 80 hertz to
08:39:29 go through.

08:39:29 And then it goes into a beamformer aspect, which
08:39:32 I'll discuss later. It also has a block called echo
08:39:36 cancellation. And this is if there is some signal that is
08:39:39 not part of what is coming from the target, such as music
08:39:44 playing or something that's repeated back through a

08:39:46 speaker, it is able to sense that and cancel that out of the -- out of the formula.

08:39:53 And there is other high pass filters that are shown in the center of the screen. There's a voice activity detector, which is a specific block that is looking for voice activity.

08:40:02 And -- and then out of the fixed beamformer, you'll notice there are seven inputs from the microphones into the beamformer. There are six outputs. So the beamformer forms six beams from the seven mic inputs.

08:40:21 These are then -- the cancellation filtration is placed in there so that the six beams that are filtered that are provided to the main beam selector, and the main beam selector selects one of those six.

08:40:36 Q. Thank you, sir.

08:40:37 And is this for a specific version of Amazon's products?

08:40:43 A. This is specifically for the Doppler.

08:40:44 Q. Thank you, sir.

08:40:46 And is there another version that you also looked at?

08:40:49 A. Yes, there's also a version called MPAF.

08:40:53 MR. RUBINO: Can we have -- please have Plaintiff's Demonstrative Slide 21?

08:41:01 Q. (By Mr. Rubino) Sir, is -- is this what you're

referring to as MPAF?
A. Yes, that is correct. This is an MPAF. It's also an audio front end, just a different type of source code that is used for this.
If you'll notice this -- this particular block diagram represents a two-by-two, so it's a microphone array. That's four microphones.
Each four of these microphones still provide those inputs into the audio front end, and, specifically, from these, it generates eight beams. The eight beams go through a filtration, and there is a -- again, a selection. So from the eight, you end up with one.
So in a similar fashion to what is done in Doppler, the Multi Platform Audio Framework also has some additional features into it, but it -- it basically operates the same.
Q. Thank you, sir.
And do you see at the top where it says Donut? Do you see that?
A. Yes, I see at the top where it says Donut.
MR. RUBINO: And so if we could have Plaintiff's Demonstrative Slide 15, please.
Q. (By Mr. Rubino) And, sir, which -- which version of the Echo is this document block diagram?
A. That -- the Donut is shown as the 3rd Generation on the

08:42:29 Dot.
08:42:29 Q. And which version do you have as a teardown next to
08:42:33 you?
08:42:33 A. The teardown for Plaintiff's Exhibit 647 is the Echo
08:42:44 Dot 3rd Generation.
08:42:44 Q. That's the same version that you have as a teardown
08:42:47 that we were just looking at on the demonstrative?
08:42:49 A. Yes, that's correct.
08:42:49 Q. In addition to the functional diagrams, did you review
08:42:53 any code, sir?
08:42:54 A. Yes, I -- I reviewed code that is applicable to both
08:42:58 Doppler and MPAF.
08:42:59 MR. RUBINO: If we could please go to Slide 22.
08:43:02 Q. (By Mr. Rubino) Sir, is this the slide you prepared?
08:43:06 A. Yes, it is.
08:43:07 Q. Can you explain what you've described here?
08:43:09 A. Yes. In looking through the various code -- source
08:43:13 code that was provided to me, the ones that I found to be
08:43:17 the most applicable were the -- the specific code modules
08:43:25 that were associated, as I show here on this Slide 22, for
08:43:28 the audio front end for Doppler.
08:43:30 Specifically, there was a modular code associated
08:43:33 with the audio front end; another code associated with the
08:43:36 fixed beamformer; also the main beam selector; a process
08:43:42 Send-in, which is the -- which is the input from the

08:43:45 microphone aspect of the module code; and
08:43:51 BeamSignalDetectorModule.
08:43:51 For the MultiPlatform, I also identified six
08:43:56 different code groups, modules that I felt were very
08:43:59 applicable to what we're going to be talking about today.
08:44:01 The first was a
08:44:07 MultiPlatformAudioFramework/Components/Module, and when I
08:44:10 used this -- this -- when the code was presented to me, as
08:44:12 I mentioned to you before Friday, it was provided in a file
08:44:16 format, and so the slash marks just represent going down
08:44:19 from the higher directory structure down to the final
08:44:25 module.
08:44:25 Also, the module for SourceLocalization;
08:44:29 FixedBeamFormer; AdaptiveBeamFormer; AEC, or
08:44:38 AcousticEchoCanceler; and FilterBank.
08:44:39 Q. Can you show us how these different code modules apply
08:44:45 to the different versions you looked at?
08:44:47 A. Yes, I can.
08:44:48 MR. RUBINO: Can we have Demonstrative Slide 23,
08:44:50 please, Mr. Thompson?
08:44:52 Q. (By Mr. Rubino) Can you explain how those code
08:44:55 versions apply to the different products, please?
08:44:57 A. Yes. The Doppler was the first code version that
08:45:00 Amazon generated for the Echo products, and they applied --
08:45:03 that particular set of code applies to the Amazon Echo 1st

08:45:12 Generation and the Amazon Echo Dot 1st Generation, so there are at least two different types of product configurations that the Doppler code applies to.

08:45:22 The MPAF multi platform, by definition it means multi platform, so it can be used across multiple different kinds of devices.

08:45:32 And you can see that the remaining products that have been accused in this case fall within the MPAF framework. And that's the Amazon Echo 2nd and 3rd Generation, the Echo Dot 2nd and 3rd Generation, and also for the Echo Dot Kids the 1st and 2nd generation, the Echo Look, 1st and 2nd Generation for the Echo Show, the Echo Spot, the 1st and 2nd Generation of Echo Plus, and the Echo Studio. They're all in conformance with the MPAF code.

08:46:04 Q. And, sir, after looking at all this code and the documentary evidence, the teardowns, and in view of the claims, what were -- what were you general conclusions?

08:46:13 A. My general conclusions were, evaluating all of the products, the teardown of the products, the source code, the documents produced in this case, deposition testimony, what I've heard from witnesses in this courtroom, is that each one of the Echo products that is identified that I've listed in Slide 23, practices Claim 1 of the -- of the '049 patent.

08:46:43 Q. Will you please look at --

08:46:51 MR. RUBINO: Will you please go to Slide 24,
08:46:54 Mr. Thompson?
08:47:04 Q. (By Mr. Rubino) Sir, can you show us how the
08:47:07 limitations of Claim 1 map to Amazon's Echo products?
08:47:11 A. I can, but because of the amount of information that's
08:47:13 included in the claim, I would prefer to go with this
08:47:17 line-by-line so that we can -- can look at the various
08:47:20 limitations within the claim without having to look at it
08:47:22 as a whole. And at the end, I'll come back and -- and put
08:47:25 it together as a whole.
08:47:27 Q. And do you have a series of demonstratives for that?
08:47:30 A. Yes, I've created some demonstratives that will take us
08:47:33 step-by-step through the claim.
08:47:35 MR. RUBINO: Can we have the next slide, please,
08:47:37 Mr. Thompson?
08:47:39 Q. (By Mr. Rubino) So, Mr. McAlexander, what are we
08:47:41 looking at here in Slide 25?
08:47:43 A. What Slide 25 is highlighting is what is known as the
08:47:48 preamble. The preamble is the introductory portion to a
08:47:51 claim. It sets the stage for what the invention is going
08:47:54 to be, which will be found in the steps that follow, but
08:48:00 this kind of sets the stage for it.
08:48:03 The preamble states: A method for enhancing a
08:48:07 target sound signal from a plurality of sound signals,
08:48:11 comprising.

08:48:11 And what will follow will be the steps that are required after this comprising preamble.
08:48:15 Q. Is there a construction you applied for this term -- this limitation?
08:48:19 A. Yes, there's a construction for the term "target sound signal" that I applied.
08:48:22 MR. RUBINO: Can we have the next slide, please, Mr. Thompson?
08:48:27 Q. (By Mr. Rubino) How did you apply this construction, sir?
08:48:29 A. The Court has construed the term "target sound signal" to be: Sound signal from a desired or target sound source.
08:48:37 And so whenever I see or observe that word "target sound signal" within the claim, this will be the definition that is applied to it.
08:48:44 Q. And at a high level, how does -- how do the Amazon products -- the accused products satisfy this portion of the claim?
08:48:55 A. Well, the target sound signal itself is -- is something that is provided from the -- the person who is voicing, for instance, the wake word. And I have a demonstrative that I think will help clearly show this.
08:49:12 MR. RUBINO: Can we go to the next slide, please, Mr. Thompson? Thank you.
08:49:16 A. So to answer your question in more detail, this is a

08:49:19 slide that comes from Plaintiff's Exhibit 1377. And, specifically, on the left side is the person who is speaking. And on the right side of this drawing is -- what is representative is the microphone array.

08:49:34 In this particular one, it's showing a -- a two microphone in -- in this arrangement, one being closer to the person who is talking and one being further away.

08:49:44 And so, clearly, if a person is speaking -- let me change color here -- if a person is speaking, that person that's speaking is going to generate a voice signal -- a target signal that is going to propagate toward this microphone array -- from the speaker to the microphone array.

08:50:11 Now, as that speaker speaks, then that voice is going to be picked up first by the microphone closest to the speaker and then, secondly, by the microphone that's further away. And so that's what -- the information is here from first to last.

08:50:34 Q. (By Mr. Rubino) And who's performing this method of Claim 1?

08:50:36 A. The method of Claim 1 is being performed by the user, the end user that is actually using this product, because when they enable the device to operate, they are providing what is necessary to do what is required here for this method of enhancing the target sound signal.

Q. And would this preamble require anything else?
A. No, it does not. The preamble, basically it's enhancing the target sound signal, and the construct of how it does that is going to be found in the later claim elements.
MR. RUBINO: May we please have Slide 28 -- Slide 29, please?
Q. (By Mr. Rubino) Sir, can you look at what's on Slide 29 and explain what this limitation is to the jury?
A. Yes. It's a rather long limitation, but I need to read it into the record. It says -- this is the first one, which I've highlighted as bracketed No. [A] -- bracketed letter [A].
And it reads: Providing a microphone array system comprising an array of sound sensors positioned in a linear, circular, or other configuration, a sound source localization unit, an adaptive beamforming unit, and a noise reduction unit, wherein said sound source localization unit, said adaptive beamforming unit, and said noise reduction unit are integrated in a digital signal processor, and wherein said sound source localization unit, said adaptive beamforming unit, and said noise reduction unit are in operative communication with said array of sound sensors.
Q. Would you like to take this claim in part, sir, this

08:52:33 limitation?
08:52:33 A. I would. Because the -- the step of providing a microphone array system comprising, and then there are a number of different aspects to that limitation after the word "comprising," so I would like to take this limitation-by-limitation.
08:52:47 MR. RUBINO: Can we have the next slide, please? Thank you, Mr. Thompson.
08:52:52 Q. (By Mr. Rubino) Sir, what have you -- what have you explained in this demonstrative?
08:52:56 A. What I am showing here in this demonstrative is you will see on the left column the entirety of this first limitation or first step of the claim. And I've highlighted in yellow a first portion, which I will address now.
08:53:09 The first portion is: Providing a microphone array system comprising an array of sound sensors positioned in a linear, circular, or other configuration.
08:53:21 And what I've shown on the right side, by way of an example, is -- when I take, for instance, the Echo device -- this is the Echo 1st Generation device that's shown on the right side of this photograph -- it is tall like a cylinder.
08:53:37 When you take it apart, in the top section underneath the -- the particular -- top of the top layer,

08:53:43 underneath that top layer, you will find this microphone array. And this microphone array is a circular board and it sits -- in fact, it's inverted, but it's right underneath the top.

08:53:59 And if you look at this microphone array, you'll see microphones. They're positioned here. Here's one, two, three, four, five, six, and one in the middle, seven.

08:54:11 So these microphones are arranged in a circular array pattern, and they communicate, as I will show later, to a digital signal processor, which is located on a sep -- separate board.

08:54:24 But in terms for this claim element, this portion of it, the microphone array, as I've shown here, has an array of sound sensors positioned in a circular configuration.

08:54:36 And I've also identified at the bottom -- I've also identified at the bottom section here other seven-microphone array systems that go back to that chart where I show all of the different products that are accused.

08:54:54 But in addition to the Echo 1st Generation, other devices that use the seven-microphone array is the -- is the Echo 2nd Generation and 3rd Generation -- 3rd Generation, the Echo Dot 1st and 2nd Generation, the Echo Plus 1st Generation and 2nd Generation, as well as the Echo

Studio. All of these use the circular -- circular seven-microphone array.

Q. Are there any other configurations that you analyzed, sir?

A. Yes, there are several other configurations.

MR. RUBINO: Can we please go to the next slide, Mr. Thompson?

Q. (By Mr. Rubino) Mr. McAlexander, are these some of the other configurations?

A. Yes. As I identified on the slide -- the slide of the accused products, there are also products that use a four-microphone array.

And what I'm showing here, by way of example, is the circular array from the Echo Dot 3rd Generation which is associated with Plaintiff's Exhibit 647 here.

In this particular type of arrangement, it's also circular, but they use four microphones instead of seven. And they're located here, one, two, three, and four.

And the other four-microphone arrays that I've identified include the Echo Dot 3rd Generation, the Echo Kids 1st and 2nd Generation, and the Echo Spot and the Echo Look.

Q. Are there any other configurations that you've looked at?

A. Yes, there's also an eight-microphone arrangement.

MR. RUBINO: Can we have the next slide, please, Mr. Thompson?

Q. (By Mr. Rubino) Mr. McAlexander, is this the eight configuration that you're referring to?

A. Yes. This one that's being shown is the eight-microphone rectangular array from the Echo Show. It also applies -- this is the Echo Show 2nd Generation. It also applies to the Echo Show 1st Generation.

And this particular arrangement uses eight microphones. In this case it's in a rectangular array, and they're located on this rectangular board. And I will circle the eight microphones located here.

Q. Sir, does this first portion of Limitation 1[A] require anything further?

A. No, it does not.

Q. How about the next portion of the limitation, can you please read that into the record?

A. The next portion, if I can have Slide 24 -- or 34. Thank you.

A sound source localization unit, an adaptive beamforming unit, and a noise reduction unit, wherein said sound source localization unit, said adaptive beamforming unit, and said noise reduction unit are integrated in a digital signal processor.

Q. Is there a construction that you applied for this term,

08:58:13 sir?
08:58:13 A. Yes, there's a construction for digital signal
08:58:17 processor.
08:58:17 MR. RUBINO: Can we please go back to Slide 33,
08:58:20 Mr. Thompson? Thank you.
08:58:21 Q. (By Mr. Rubino) Is this the construction you were
08:58:23 referring to, sir?
08:58:24 A. Yes, it is. The construction for digital signal
08:58:29 processor is: Microprocessor that is specialized for
08:58:36 mathematical processing of digital signals.
08:58:38 So as I stated before, any time in the claim where
08:58:42 the word "digital signal processor" occurs, this is the
08:58:44 definition that applies.
08:58:46 MR. RUBINO: Can we have Slide 34 again, please,
08:58:49 Mr. Thompson? Thank you.
08:58:51 Q. (By Mr. Rubino) Sir, can you explain how you found
08:58:53 that in the products?
08:58:54 A. Yes. The number of the documents I reviewed, as well
08:58:59 as my own personal teardown of the accused products,
08:59:02 identified the microphone array that was located on the
08:59:04 board that I've already identified to the Court. And then
08:59:07 I looked at what the output of the array is -- is
08:59:11 communicating to. And in each case, the array communicates
08:59:15 to a digital signal processor.
08:59:20 What I have identified here on the right side of

Slide 34 are the digital signal processors that are found in the various products.

As an example, the Echo 1st Generation and the Echo Dot 1st Generation use a Texas Instruments digital signal processing chip labeled as the DM3735.

For the Echo 2nd Generation and the Echo Dot 2nd Generation, the Echo Spot, and the Echo Show 5, they use a digital signal processor that is identified as a MediaTek MT8163 device.

In the third line, I've identified the MediaTek MT8516 digital signal processor that is used for the Echo Plus 1st and 2nd Generation and the Echo Studio.

In the fourth row, I've identified for the Echo Dot 3rd Generation and the Echo Dot Kids Versions 1st and 2nd Generation the MediaTek MT7658 device.

And, lastly, at the bottom, for the Echo Look in both the 1st and 2nd Generation of the Echo Show, I have identified the digital signal processor that is known as the Intel Cherry Trail T3. It's also identified as the Atom X5-Z850 device.

Q. Sir, do you have any other sources that you've looked at to confirm that these are digital signal processors?

A. Yes. For each one of these devices, I've also identified what's called manufacturer specifications. The -- the producer of these devices provides to the public

specification sheets that's sometimes referred to as data sheets, but it provides the specifics of the device, how it's configured, how it's packaged, how it's -- what kind of timing signals you provide to it, where you connect to it for input and output. So it basically just describes the overall operation of the device.

MR. RUBINO: Can we have Slide 35, please, Mr. Thompson? Thank you.

Q. (By Mr. Rubino) Can you explain what -- what this is, sir?

A. As I had described earlier, the MT8516 is one -- by way of example, this is the digital signal processing device that's used for the Echo Plus 1st and 2nd Generation and the Echo Show. And this is a copy of one of the pages out of the MT8516 sheet.

And if you'll notice, this is also from MediaTek, so they're the ones who produced the device. And the MT8516 is defined as a highly integrated, application processing platform.

And if you will notice, in the second paragraph, it states that: MediaTek's unique PowerAQ tool provides an easy GUI -- that's graphical user interface -- for signal -- for signal flow design and audio parameter tuning, removing the need for an additional DSP.

So they have embedded digital signal processing

09:02:35 capability on this device itself, so there's no need for a
09:02:39 separate DSP chip.
09:02:41 Q. Sir, have you ever heard of something called Neon?
09:02:45 A. Yes, I have.
09:02:46 Q. In the context of digital signal processors?
09:02:48 A. Yes, I have. In fact, for the MT8516 device, if you'll
09:02:53 notice in the third row -- the third paragraph, it states
09:02:57 that: The MT8516 integrates a quad-core, 64-bit ARM
09:03:07 Cortex-A.
09:03:08 Cortex-A, by definition according to ARM, includes
09:03:12 the Neon.
09:03:13 MR. RUBINO: Can we have Slide 36, please,
09:03:15 Mr. Thompson? Thank you.
09:03:17 Q. (By Mr. Rubino) Sir, can you explain what -- what that
09:03:22 Neon is in the context of ARM?
09:03:25 A. Yes. I had mentioned that the ARM Cortex-A series,
09:03:30 which was the previous slide, includes the Neon device.
09:03:37 And the Neon device shows to be an advanced single
09:03:41 instruction multiple data architecture extension for the
09:03:46 ARM Cortex-A series.
09:03:48 And so this, according to the second sentence, the
09:03:50 ARM Neon accelerates audio, encoding, and so forth. Neon
09:03:56 also accelerates signal processing algorithms and functions
09:04:02 to speed up applications such as audio processing.
09:04:06 Q. Sir, how does -- how does Amazon use this Neon

09:04:12 architecture?
09:04:12 A. The Neon architecture is embedded in the Cortex-A,
09:04:22 which is part of the ARM core. ARM core is basically the
09:04:27 processing -- the high-end processing structure that's
09:04:28 built into this device.
09:04:29 And so Amazon uses this because the information
09:04:33 that is provided to it from the microphones, the samples
09:04:38 that are taken from the microphone for the voice that is
09:04:42 spoken, are provided as inputs to this digital signal
09:04:45 processing chip. And it is the Neon aspect of the ARM
09:04:50 core, that actually provides the handling of this
09:04:56 information.
09:04:56 Q. And does this limitation require anything else, sir,
09:04:58 this portion of the limitation?
09:05:00 A. No, sir, it does not.
09:05:06 MR. RUBINO: May we please proceed to Slide 37,
09:05:09 Mr. Thompson?
09:05:10 Q. (By Mr. Rubino) And Mr. McAlexander, can you read the
09:05:13 final portion of Limitation 1[A]?
09:05:17 A. Yes. The final portion is: And wherein said sound
09:05:21 source localization unit, said adaptive beamforming unit,
09:05:22 and said noise reduction unit are in operative
09:05:24 communication with said array of sound sensors.
09:05:26 And what I'm showing here by way of example is the
09:05:32 four-microphone array board. This would apply to the

seven-microphone array or the eight-microphone rectangular, any one of those. But I will describe in more detail later how the sound source localization unit, the adaptive beamforming unit, and the noise -- noise reduction unit operate within this digital signal processor.

But for purposes of this particular claim limitation, it just says that those particular operative structures -- the code that's operating and executing the digital signal processor located in the middle of this board is in operative communication with each of the microphones.

And the -- if you trace the signals themselves on the -- on this particular board, you will find that the -- the signals that are propagated from the microphones do, in fact, communicate to the digital signal processor, and then they're handled by the code that's executed on that digital signal processor.

Q. Sir, does this portion of the limitation require anything further?

A. No, it does not.

Q. And so this limitation starts with the word "providing." Do you see that?

A. Yes, I see that.

Q. Now, who's doing the providing here?

A. The end user who is using the device is doing the

providing.

MR. RUBINO: Can we go to Slide 38, please, Mr. Thompson?

Q. (By Mr. Rubino) Can you explain that to the jury, please?

A. Yes. As informed -- as the public is informed by Amazon, this is one of the documents that shows this. This is from Plaintiff's Exhibit 1372.

And Amazon is instructing the public, the end user, that this is how you set up your Echo. Basically, it says plug it in. And so plug it in, run the app, everything is ready to go.

And so the only thing that's required is to speak the wake word. When you speak the wake word, that's exactly where the invention that -- of the '049 takes place is that -- is there a response to that wake word in the inventive way.

So the person who is installing and using this system has been informed by Amazon to turn it on and use it in a way that they specify, and when they do so, they are provide -- they are doing the providing step.

Q. Sir, in the context of this first claim limitation, you said you looked at technical specifications; is that right?

A. Yes, that's correct.

MR. RUBINO: And if we could have Plaintiff's 441,

09:08:28 please. 441, please. Thank you, Mr. Thompson.
09:08:31 Q. (By Mr. Rubino) What kind of a document is this, sir?
09:08:36 A. This is -- it's identified as Project Biscuit, and it's
09:08:42 a development commitment document, so it would be an
09:08:45 internal document that was generated with regard to the
09:08:50 Biscuit project.
09:08:50 Q. And is this one of those types of technical
09:08:53 specifications you looked at?
09:08:54 A. Yes, this is one of the tech -- technical
09:08:57 specifications that were provided. And this particular
09:08:59 document comes from Amazon Lab126, which was identified as
09:09:06 part of their research that's part of Amazon.
09:09:10 MR. RUBINO: Can we please have PTX-321? Thank
09:09:14 you, Mr. Thompson.
09:09:20 Q. (By Mr. Rubino) And, additionally, what type of
09:09:22 document is this, sir? Is this one of those documents as
09:09:26 well?
09:09:26 A. This is a document identified for the Bishop -- by the
09:09:30 way, the Biscuit was the 2nd Generation Echo Dot. The
09:09:31 Bishop that we're referring to here is the 2nd Generation
09:09:36 Echo Show.
09:09:37 Q. Thank you, sir.
09:09:39 Ans so for this claim limitation --
09:09:41 MR. RUBINO: If we can go back to the
09:09:43 demonstratives.

Q. (By Mr. Rubino) -- does it require anything further, sir?
A. No, it does not.
MR. RUBINO: Can we have Slide 39, please? May we please go to Slide 40, Mr. Thompson?
Q. (By Mr. Rubino) Mr. McAlexander, can you please read Limitation [B] into the record?
A. Receiving said sound signals from a plurality of disparate sound sources by said sound sensors, wherein said received sound signals comprise said target sound signal from a target sound source among said disparate sound sources and ambient noise signals.
Q. And what does this limitation require, sir?
A. It requires that the -- basically, the way it operates, the microphone array is going to receive sound signals, and we've already identified the fact that the products have a microphone array. So this is the receiving aspect. It's receiving sound signals that are generated.
And it states that the sound signals are going to be a combination of the target sound signal, in other words, the one you're really looking for. If the person speaking the wake word, that's the one you want to discriminate and find out of all the signals that are coming in.
But of the signals that are being in -- that are

coming in or being received by the Amazon device, there are going to be other sound sources, such as I described Friday. There's going to be ambient sound sources, sound from, for instance, the 60-cycle hum from the electrical lights that are on. We had a quiet time on Friday where you could literally hear what the lights were doing.

So there's going to be ambient noise in the environment. Air conditioning provides an ambient noise.

And there are also disparate sound sources. These are -- can be represented as other major sound sources. For instance, one person speaking the wake word but may be standing around in the same room as another person talking. Well, that's another verbal sound, but it's not the target sound signal.

So there can be disparate sources or ambient sources, and the key for this invention is how to detect and pull out the specific target sound signal.

Q. Is there a construction in this term, sir?

MR. RUBINO: If we could have Slide 26, please.

A. It is the same one word that we talked about in the preamble, the target sound signal. Again, any time you see the word target sound -- target sound signal, the definition for that term is: Sound signal from a desired or target sound source.

Such as an example that I'm using, if a person who

is infringing this claim is speaking the wake word, the speaking of the wake word, the wake word is the -- is the target sound signal.

MR. RUBINO: Can we please go back to Slide 41, Mr. Thompson? Thank you.

Q. (By Mr. Rubino) Sir, can you -- can you show us any evidence of when this limitation is met?

A. Yes. This is -- I'll call it an indication that the claim is being practiced. For instance, if -- if I look at any one of these Amazon products, as I'm showing here, around the top where the -- remember, at the top of the cylinder, I said underneath it is the microphone, but if you look at the top, there's also some holes or openings around that circle.

And there's also a ring of lights. In fact, there's 12 LEDs in some of these models that are in a circular array in the same general vicinity as the microphones.

And so when you speak and the Amazon device detects speech, detects something that is happening, it will immediately light this ring up. It will -- it will light it up into a darker blue color. And when it immediately, very quickly discerns the direction of the targeted sound signal, for instance, the wake word, when it detects that wake word and the direction it's coming from,

09:14:07 that particular portion of the light ring will turn to a
09:14:11 lighter blue.

09:14:12 So if you'll notice in this example that I'm
09:14:14 showing, you'll see the darker color of the ring, but
09:14:17 you'll also see on the left side -- you will also see on
09:14:26 the left side here the blue has turned to a lighter blue.

09:14:36 And so that is an indication of the fact that the
09:14:39 system has made the decision -- has gone through the
09:14:41 process to determine the direction from which that wake
09:14:45 word is coming and has turned the light on in the direction
09:14:47 of the user.

09:14:48 So if you're actually using this device, you will
09:14:51 see it when you speak the wake word, the Amazon light ring
09:14:55 will light up, and then it will turn light blue in the
09:14:57 direction that you're talking from.

09:15:01 Q. Sir, did you look at any technical specifications or
09:15:04 source code for this limitation?

09:15:05 A. Yes, I did.

09:15:07 Q. And did you find this limitation there, as well?

09:15:10 A. Yes, I did.

09:15:12 MR. RUBINO: Can we have PTX-12, please?

09:15:15 Q. (By Mr. Rubino) Mr. McAlexander, is this one of those
09:15:24 documents that you looked at in rendering your opinions?

09:15:26 A. Yes. PTX-12 is another document produced by Amazon's
09:15:30 Lab126, and this particular document is identified as the

audio front end, AFE, architecture for the Biscuit, the Sonar, and the Knight.

MR. RUBINO: Can we go to Page 6, please, Mr. Thompson?

Q. (By Mr. Rubino) Mr. McAlexander, does this figure show anything related to your analysis for this claim limitation?

A. Yes. At the top of this page, which is -- I did not catch the page number -- Page 6 of 12, which is Amazon's Bates Stamp No. 6283 at the bottom right. So if we look at Figure 1, it's a system block diagram for the audio front end for the Doppler, which is the one that applies to the original Echo and the Echo Dot products.

And we'll see on the -- on the upper left side the -- the seven-microphone array. And the seven-microphone array picks up that signal and picks up the signals and provides it in an instruction called Send-in in the code, and that basically provides the input to the system. So this highlights receiving sound signals according to what the claim requires.

Q. Did you look at any documents for the MPAF, as well?

A. Yes, I did.

MR. RUBINO: Can we have PTX-115, please? Thank you, Mr. Thompson. Can we go to Page 2, please? Sorry, Page -- Page 3.

09:17:22 Q. (By Mr. Rubino) Sir, does this document show you
09:17:24 anything about this limitation, Limitation [B] for the MPAF
09:17:32 products?
09:17:32 A. Yes. This particular one is the -- the audio front end
09:17:41 as shown here for the Sonar, which is the next generation.
09:17:44 It's an MPAF-type system. And just as in the Doppler,
09:17:48 you'll see, once again, the microphone array providing an
09:17:52 input using Send-in into the system itself.
09:17:56 And so, once again, this document at this
09:18:03 diagrammatic level shows that there is, in fact, a
09:18:07 receiving component to the accused products.
09:18:09 Q. Thank you, sir.
09:18:09 And does this limitation, Limitation 1[B] require
09:18:16 anything further?
09:18:17 A. No, it does not.
09:18:18 MR. RUBINO: Can we turn to Slide 42, please,
09:18:22 Mr. Thompson?
09:18:22 Q. (By Mr. Rubino) So next is Limitation 1[C]. Can you
09:18:28 please read that limitation into the record, sir?
09:18:30 A. Yes. Limitation 1[C] states: Determining a delay
09:18:38 between each of said sound sensors and an origin of said
09:18:42 array of said sound sensors as a function of distance
09:18:44 between each of said sound sensors and said origin, a
09:18:48 predefined angle between each of said sound sensors and a
09:18:53 reference axis, and an azimuth angle between said reference

09:18:58 axis and said target sound signal, when said target sound
09:19:02 source that emits said target sound signal is in a
09:19:06 two-dimensional plane, wherein said delay is represented in
09:19:10 terms of a number of samples -- excuse me -- is represented
09:19:14 in terms of number of samples, and wherein said
09:19:16 determination of said delay enables beamforming for said
09:19:21 array of sound sensors in a plurality of configurations.
09:19:25 Q. Sir, at a high level, how is that claim limitation met?
09:19:30 A. The determine -- the determining a delay steps is for
09:19:34 purpose of enabling beamforming. And so there's
09:19:37 beamforming algorithms that are provided within each one of
09:19:39 the Doppler and MPAF code that, when executed by the
09:19:45 digital signal processor, performs this step of determine a
09:19:50 delay. And it basically is -- is following a -- a -- a
09:19:55 filter-and-sum technique for forming -- beamforming. And
09:19:59 I'll show you in more detail how that works.
09:20:03 Q. Sir --
09:20:04 MR. RUBINO: If we could go to Slide -- to
09:20:07 PTX-1377, please, at Page 60. Next page, please. Thank
09:20:19 you, Mr. Thompson.
09:20:21 Q. (By Mr. Rubino) Mr. McAlexander, is this one of the
09:20:23 slides we looked at earlier?
09:20:24 A. It is -- that is correct. It is -- it is showing, in
09:20:28 this, case six different microphones in array.
09:20:32 Q. Does this depiction indicate anything to you about the

09:20:36 determination of a delay limitation?

09:20:38 A. It does because the -- this is the indication, as I mentioned before, is that when a sound is emanating from the person on the left of this diagram, that sound is going to be basically broadcast in a lot of different directions so that people in different parts of the room can hear.

09:21:05 But in terms of where the Amazon device is located, it's going to pick up the sound signal that is targeted from that sound source or the person in the direction that it's -- that it's coming into the Amazon device.

09:21:16 And so the way it's diagrammed here is certainly the -- the microphone that's identified as first is going to be the first microphone that picks up the signal. The signal is a wave that's being propagated from the person in the direction of the Amazon device.

09:21:35 So as it is being received by the Amazon device, the microphone closest is going to pick up the signal first, and the microphone that's further away is going to pick it up later.

09:21:45 And then when you have microphones that are positioned on a slightly different azimuth, an in-coming angle, they're going to be picking them up differentially, too. So each microphone is going to be picking up the signal at a different time. It's a delay.

09:22:02 Q. Thank you, sir.
09:22:04 And earlier did you mention beamforming?
09:22:05 A. Yes, I did.
09:22:06 Q. Can you read that first line of the slide?
09:22:08 A. Yes, it -- this slide is directed to the audio algorithms. And it says: Beamforming is locate the source of speech and pick it out of the background noise.
09:22:19 So this is receiving the -- the targeted signal, as well as disparate signals, signals in ambient noise. All of that is being received.
09:22:28 And now the beamforming is to locate the source of the speech and pick it out of the background noise.
09:22:34 Q. Thank you.
09:22:35 And where is this slide from, sir?
09:22:37 A. This is from Plaintiff's Exhibit 1377, which is a presentation.
09:22:47 Q. What is your understanding about that, sir? Who is it from?
09:22:50 A. Well, this is an Amazon presentation, and the title of it was the AWS re:INVENT, Integrate Alexa voice technology into your product with the Alexa Voice Service. And that's the first page of the -- Plaintiff's Exhibit 1377.
09:23:11 And the source, to answer your question, it shows copyright 2017 from Amazon Web Services, and it's AWS.
09:23:22 Q. Can you explain this claim at a more granular level?

09:23:26 There were a lot of words there in this limitation. Is it
09:23:29 possible for you to break it down?
09:23:45 A. Well, as I did in the previous claim element, there are
09:23:50 a number of different limitations that are shown here. So
09:23:53 to answer your question, yes, we will look at it in a
09:23:57 little bit more of a granular level, and basically look at
09:24:00 the first section, and then we'll do it in stages and step
09:24:04 through this particular claim element.
09:24:05 MR. RUBINO: Can we go to Slide 46, please,
09:24:08 Mr. Thompson? Thank you.
09:24:10 Q. (By Mr. Rubino) Sir, how have you -- how have you
09:24:12 broken it down here, sir?
09:24:14 A. Well, the first section that I've identified of this
09:24:17 claim step is: Determining a delay between each of said
09:24:22 sound sensors and an origin of said array of said sound
09:24:26 sensors as a function of distance between each of said
09:24:28 sound sensors and said origin.
09:24:30 Now, clearly, for each one of the Amazon products,
09:24:34 the array that has been chosen embedded in that product in
09:24:38 terms of the microphone array is fixed. It's already
09:24:41 pre -- pre-identified, know exactly where the location is
09:24:45 of each of the microphones, and they do -- they themselves
09:24:48 physically do not move from that orientation.
09:24:50 So the -- the structure in terms of the center
09:24:53 between the microphone array, as well as the position in a

coordinate system -- in an angle from that origin, is fixed.

Q. And have you found evidence of this in the source code, sir?

A. Yes, I have.

MR. RUBINO: Can we please turn to Plaintiff's -- PTX-386?

Q. (By Mr. Rubino) Sir, what is this document that we're looking at?

A. This is a document that was authored by Amit Chhetri, and this is also out of Amazon's Lab126. And, specifically, it's directed to computing the two-dimensional beampatterns.

And we'll see that, in this document, that there's a particular algorithm -- formula that is used, and states that the beampattern is a function of elevation angle, azimuth angle, and frequency.

And so in this particular arrangement, it says it's for a given -- sentence right above it says: For a given beamformer designed with a microphone array, compute the two-dimensional beampatterns.

And so this is a construct by which they take the instantiation of a particular microphone array organized in a certain fashion, arranged in a certain architecture at a distance from a center, and create what I will call the

09:26:30 weighting factors that are associated with -- with that particular kind of a structure.

09:26:34 So what -- what is done by Amazon is that they take each structure, they model it, and -- and it's called -- modelled in MATLAB or COMSOL, two different kinds of programs.

09:26:46 But within the modelling, what they do is they take a specific designed architecture with a physical arrangement of the microphones, and then they simulate signals coming in from different azimuths and different elevations.

09:27:02 And -- and then they -- from that, they determine how that matrix of microphones will be able to best discriminate the source of the sound, and they provide weighting factors.

09:27:13 And -- and that goes into the -- the initial construct of how the beams will be formed once they are instantiated in the accused device.

09:27:24 Q. Now, sir, when you say the beams will be formed, where are those beams eventually formed?

09:27:28 A. They're eventually formed when the user uses the system. The beams are formed at that time. And they're formed based upon weighting factors that were pre-determined in the laboratory from a simulation model, which simulates all the different kind of combination of

09:27:44 signals that can come in so that the unit in operation will know what -- what to pull in in order to create those beams.
09:27:51 Q. Sir, you mentioned distances between microphones and -- in this limitation. Did you find that, as well?
09:28:01 A. Yes. The microphone array, they're computing elevation angle, azimuth angle, and frequency. But the weighting factors include the understanding of how that arrangement of the architecture is because one of the key factors is designed with a microphone array. So you have to define the microphone array first.
09:28:24 Q. Thank you, sir.
09:28:25 In the top line of this document, do you see where it says "function"?
09:28:29 A. Yes, I do.
09:28:31 Q. And the first input says micLocM. Do you know what that means?
09:28:40 A. Well, the first location is micLocationM. And this is where in this simulation model you specify the architecture, the arrangement of the microphones. So this is the location of the microphones. And that's one of the inputs into this algorithm.
09:28:57 MR. RUBINO: Can we please have Plaintiff's 358, Mr. Thompson?
09:29:08 Q. (By Mr. Rubino) Sir, do you know what this document

09:29:10 is? Have you looked at this one?
09:29:12 A. Yes, I have. PTX-358 is directed to beamforming
09:29:20 specifically for MPAF. MPAF adds an additional element to
09:29:23 it that is not found in the Doppler code.
09:29:24 But MPAF does what's called subband. It's a --
09:29:31 it's -- it's taking the frequency signals that are coming
09:29:34 in, and it divides them.
09:29:36 For instance, if you're -- you're speaking -- and,
09:29:38 typically, speaking is certainly above 80 hertz. And,
09:29:44 typically, it's around -- up to 3500; maybe at the most
09:29:47 8,000. We can hear from like, generally speaking, 20 to
09:29:51 20,000 hertz. But, typically, speaking ranges is in the
09:29:56 hundreds to the thousand -- low thousands number of
09:29:58 frequency.
09:29:58 And so it's taking this frequency and dividing it
09:30:01 into certain bands. And then when it internally processes
09:30:06 this with the high math algorithms, it's doing it on a
09:30:10 subband basis. It's quicker and it's more refined, and it
09:30:14 gives a more exact result. And that's done in MPAF but not
09:30:18 in Doppler.
09:30:18 MR. RUBINO: Can we zoom out, Mr. Thompson? Thank
09:30:21 you. Can we turn to the next page of the document,
09:30:28 Mr. Thompson? Thank you.
09:30:30 Q. (By Mr. Rubino) Mr. McAlexander, what is this page of
09:30:33 the document showing?

A. This page of the document is showing a function -- and if we can enlarge the -- yes, that's correct.

Notice that in the previous slide, I had talked about microphone location, micLocM, and so -- performing this function with the index based upon the microphone array configuration, it also includes the number of mics.

And if you'll notice -- for instance, the way in which this operates is that this particular code illustration that I've identified -- and there are others, so this is just one of the illustrative areas of the code -- is that if you have a two-microphone array, it shows the -- the distance X, Y, and Z coordinates, so you can see it's located at no elevation. And there's the X/Y components.

So from an origin point, it shows the displacement in one or more directions of a particular microphone array.

Q. And so with regard to that first portion of the limitation --

MR. RUBINO: If we can go back to Slide 47, Mr. Thompson -- sorry, 46. Thank you.

Q. (By Mr. Rubino) So, Mr. McAlexander, with regard to that first portion of the limitation, does it require anything else?

A. No, sir, that -- that completes what's necessary to show the dis -- comprising a function of the distance

09:31:57 between the sensors and the origin.

09:32:02 Q. Now, have you addressed the second portion of this limitation?

09:32:06 A. Yes, I have.

09:32:07 Q. And what meets this second portion of the limitation?

09:32:10 A. Well, let me read into the record a second portion. It states: A predefined angle between each of said sound sensors and a reference axis, and an azimuth angle between said reference axis and said target sound signal.

09:32:24 There's two aspects that are provided here. One is the angle that is between the sound sensors.

09:32:29 And, second, it's terms of the azimuth angle, and that's the incoming sound signal.

09:32:37 Q. Did you find that in the code somewhere, sir?

09:32:39 A. Yes, I did.

09:32:40 MR. RUBINO: Can we please turn back to PTX-386, Mr. Thompson?

09:32:46 Q. (By Mr. Rubino) Can you explain that further, Mr. McAlexander?

09:32:50 A. Yes.

09:32:51 If we can enlarge this section. No, not quite. Back up. There we go.

09:33:02 Once again, the first function I talked about was the mic location. And here in the micTune and the micLook are two other parts in this function. And the left is the

09:33:21 steering. Which direction is the target signal coming from so I can identify the -- the look direction, or which way do I look?

And the tune, this -- this also looks at the differentiation of all the signals coming in. It's specifically creating the beampatterns that are necessary for selecting which one of the beams to use. And that's done in micTune.

09:33:44 Q. Thank you, sir.

And does this portion of the limitation require anything else?

09:33:49 A. No, because the beam -- as I stated here before, the beampattern is the function of elevation angle, azimuth angle. And so all of that is covered and -- and meets the limitation of the angle between the sound sensors and the azimuth of the target sound signal that's coming out.

09:34:07 MR. RUBINO: If we can please have Slide 48, Mr. Thompson. Thank you.

09:34:11 Q. (By Mr. Rubino) What about the next portion of this claim term, sir?

09:34:13 A. The next portion reads: When said target sound source that emits said target sound signal is in a two-dimensional plane.

So the additional limitation here is the source signal coming in from a two-dimensional plane.

09:34:26 Q. Is there a construction of this term that you've applied?
09:34:29 A. Yes, there is.
09:34:30 MR. RUBINO: Can we have the next slide, Mr. Thompson?
09:34:33 Q. (By Mr. Rubino) Is this that construction that you applied, sir?
09:34:35 A. Yes. The term that when said target sound source that emits said target sound signal is in a two-dimensional plane, the definition that has been provided for us to use is: When said target sound source that emits said target sound signal is treated as if it is in the same two-dimensional plane as the sound sensors. And so this would be the definition that would be put in, to define that particular claim term.
09:35:02 Q. And have you found that term, sir?
09:35:03 A. Yes, I have.
09:35:04 Q. Have you found that in any code?
09:35:12 A. I found that in the code, basically, the Doppler versus in the MPAF.
09:35:16 MR. RUBINO: Can we please go back to PTX-386, please?
09:35:19 Q. (By Mr. Rubino) Where have you found it in the code, sir?
09:35:23 A. Well, this requires to treat the signal as if it is in

09:35:30 the same two-dimensional plane as the sound sensors. In
09:35:34 this particular illustrative portion of the MATLAB code
09:35:39 is -- circles don't work when the screen moves.
09:35:45 You'll notice at the top it's two-dimensional
09:35:50 beampatterns.
09:35:51 Q. So does this limitation require anything else, sir?
09:35:54 A. No, it does not.
09:35:57 Q. Now, we've taken a look at this code file. You said
09:36:01 the word exemplary -- did you say exemplary?
09:36:04 A. I think I said illustrative.
09:36:07 Q. Illustrative, excuse me.
09:36:07 A. Yeah.
09:36:08 Q. What did you mean by illustrative, sir?
09:36:10 A. There are other -- there are other portions or modules
09:36:13 of code that are embedded within all of the code that I
09:36:16 looked at. This is just one of them. And very -- and one
09:36:19 of the ones that I mentioned that I looked at are very
09:36:22 similar, using the same formulation as what's shown on this
09:36:26 page.
09:36:27 Q. And does this limitation require anything further?
09:36:29 A. No, it does not.
09:36:31 MR. RUBINO: Can we please go to Slide 50,
09:36:33 Mr. Thompson? Thank you.
09:36:34 Q. (By Mr. Rubino) What about the next portion of this
09:36:45 claim limitation?

09:36:46 A. The next portion of this reads: Wherein said delay is represented in terms of number of samples.
09:36:52 Q. Did you find that limitation, sir?
09:36:53 A. Yes, I did.
09:36:54 Q. Where did you find that limitation of the products?
09:36:57 A. I found it in the technical documents that were produced in this case, including the specifications, the block diagrams, as well as in the code.
09:37:04 Q. Where did you find that in the code, sir?
09:37:06 A. I found it in the beamformer -- beamforming code sections.
09:37:13 MR. RUBINO: Can we please put up Plaintiff's 1378?
09:37:23 Your Honor, for this exhibit, if I may use the ELMO.
09:37:53 Q. (By Mr. Rubino) Sir, is this PTX-1378?
09:37:56 A. Yes, it is.
09:37:56 Q. And can you just give a high-level overview of what it is that you compiled in this document?
09:38:01 A. This document, PTX-1378, as shown in the bottom left corner, almost visible, it says audioFrontEnd.cpp, and this is C++, this is the type of source code that is used.
09:38:20 And in this audio front end, it basically -- this, among a number of other portions of the code, talk -- discuss the signals that are propagated and received and

09:38:32 how it represents the delay in terms of number of samples.
09:38:37 And I -- I might mention, before we look at this,
09:38:39 is that on previous slides, I had shown where the
09:38:43 microphone array receives the signals and it propagates
09:38:47 that into the device, into the DSP, the digital signal
09:38:53 processing portion of the device.
09:38:54 And each one of those signals that I identified as
09:38:58 Send-in comes in at 16,000 bits per second. So it's
09:39:02 sampled at 16,000 bits -- 16,000 samples per second. And
09:39:08 then that input then is provided further in the code for
09:39:11 processing.
09:39:12 Q. Sir, I've put up a document marked PTX-1378-03.
09:39:22 Could you explain what this document is, at a high
09:39:25 level?
09:39:25 A. Yes. This is -- did you say 03?
09:39:38 Q. Yes, sir.
09:39:39 A. Okay. All right. This document is also part of the
09:39:43 Doppler code and is from the module. It's the super
09:39:50 directive beamformer code, sdb -- sbd.c++ -- cpp.
09:40:00 And in this particular one, this invokes the
09:40:02 processing of the incoming information, and it does that --
09:40:08 if you'll notice, it says, after setting the microphone
09:40:14 input data properly into the buffers, so this is the input
09:40:17 data that's sampled at 16,000 bits per second. That
09:40:21 sampling is provided and stored in this particular area.

It says: Input properly into the buffers. Buffers is a storage of those input -- that input information.

And it then states: Perform beamforming. Now, the beamforming that is performed is done using what is called a -- it's a high-level math called Fast Fourier Transform.

And what Fast Fourier Transform does is it takes the input samples and converts it from the time domain, because the samples are coming in on a time basis as they're received by the microphone. And it's a high-end math that converts this, transforms it into the frequency domain, and does all the analyticals in the frequency domain of the samples and creates sample output as a result of that.

And then once it finishes that beamforming process, it goes through what's called a synthesizer, and it inverts it back to the time domain.

Q. Sir, can you point us to some samples in this document where you see that Fast Fourier Transform?

A. Yes. Every time you see the term FFT, for instance here, I mentioned to you the descriptive language, this is the description that's actually written in there. The actual function is performed here. It's called memory set. And this is where the buffers are set. It uses the term

09:41:46 FFT, which is Fast Fourier Transform.
09:41:51 Q. Thank you, sir.
09:41:52 For this limitation, is anything else required for
09:41:55 this portion of the limitation?
09:41:57 A. No, it is not.
09:41:58 Q. What about the final portion of the limitation?
09:42:00 MR. RUBINO: If we could go to Slide 51, please.
09:42:17 A. By the way, I did want to mention when we go to this
09:42:21 one, that everything is met. What I'm showing is code for
09:42:24 the Doppler, and it will be similar for the MPAF.
09:42:28 The last limitation portion of this is: Wherein
09:42:31 said determination of said delay enables beamforming for
09:42:34 said array of sound sensors in a plurality of
09:42:40 configuration.
09:42:40 This is basically the end product, the end result
09:42:43 of this delay -- of the determining a delay step.
09:42:46 Determining a delay step is -- is using samples for the
09:42:50 purpose of enabling beamforming. Beamforming is what the
09:42:54 result of this step is.
09:43:04 Q. (By Mr. Rubino) And how is that -- how is that step
09:43:07 done by the Amazon products, sir?
09:43:08 A. Well, it is -- I'll show how it's done, but I do think
09:43:12 it's important that there is a particular claim term that
09:43:14 needs to be covered here, and that's plurality of
09:43:19 configurations. We have been provided a definition for

09:43:22 that.

09:43:22 So I want to make sure that the jury understands
09:43:25 this term "for said array of sound sensors in a plurality
09:43:29 of configurations requires the definition": For said array
09:43:32 of sound sensors in a plurality of geometric layouts of the
09:43:40 sound sensors.

09:43:41 Q. And how is that met?

09:43:48 A. The -- once the samples that are received by the
09:43:52 microphones are sent in using the Send-in statement into
09:43:58 the digital signal processor, they are then processed in
09:44:03 the frequency domain using Fast Fourier Transform. It
09:44:08 samples in, evaluates that. And the output of that is
09:44:16 beamforming. So it identifies and determines those
09:44:18 particular set of beams.

09:44:21 For instance, if you have seven microphone in a
09:44:24 circular array, it will actually instantiate six different
09:44:27 beams as a result of that. It's basically the combination
09:44:29 of those seven will result in six beams. And so that is
09:44:33 done by the beamforming algorithms that I showed.

09:44:38 Q. And how does that work across the products, sir?

09:44:41 A. That works across the products in terms of both the 1st
09:44:45 Generation device for the Dot and the Echo using the
09:44:52 Doppler code and all the other accused products using the
09:44:55 MPAF code, that the result is that there is a set of beams
09:44:59 that are determined and provided as a result of that.

09:45:04 And then from that we'll see later steps where the one that is the most prominent one is selected.

09:45:10 MR. RUBINO: If we could turn back to Slide 51, please.

09:45:15 Q. (By Mr. Rubino) And how about for the MPAF version, sir?

09:45:18 A. The same is true. The determination of that delay is done by beamforming. The beamforming part of that algorithm is done -- is done what's called the filter-and-sum technique. It's a standard technique that is employed, is well-known.

09:45:34 And in that particular domain, using the filter-and-sum technique, the delay -- everything that's done for the delay is built into that code. And that's including the -- the arrangement of the mics and everything for the determination of the azimuth angle and the -- and the conclusion as to how to weight the beams in accordance with that direction of the sound signal coming in. It's all done in that code.

09:46:01 Q. And is that in a plurality of configurations?

09:46:02 A. Yes. The reason it's done in a plurality of configurations is because, for instance, Doppler is -- the Doppler can be utilized across two different products. It's a plurality of configuration.

09:46:16 The MPAF can be provided across all the other

09:46:18 products; again, a plurality of configurations.
09:46:20 Q. Do you have a high-level summary of this claim
09:46:27 limitation that you've put together as a demonstrative?
09:46:29 A. Yes, I do.
09:46:31 MR. RUBINO: Can we please turn to Slide 53?
09:46:35 Q. (By Mr. Rubino) Can you explain your high level here,
09:46:37 sir?
09:46:37 A. The high level here is attempting to -- in an
09:46:43 illustration diagram to just show -- show you where we are
09:46:49 at this point based upon the analysis of the claim as
09:46:51 applied to the accused devices.
09:46:56 When you have an incoming signal, that incoming
09:46:59 signal from the person on the right side speaking the wake
09:47:04 word is going to be generating a signal -- a target signal
09:47:11 that's coming into the device.
09:47:12 It's going to be picked up by the microphone
09:47:14 speakers, which are located in the center -- central ring
09:47:19 just below where you see these openings.
09:47:22 And, as I mentioned before, the input that goes in
09:47:27 the microphones then is provided to the DSP device in the
09:47:32 middle, digital signal processing device.
09:47:35 And so these incoming signals will be provided on
09:47:40 a sample basis, will be processed using the Fast Fourier
09:47:47 Transform. And from that, there will be a determination in
09:47:50 that Fast Fourier Transform, using the filter-and-sum

09:47:53 technique, will then determine the beams. So the beams will be determined based upon the delay that's calculated in that particular summation aspect.

09:48:03 And then I've noted here on the left side that the simulation models have provided some beamforming weights, and that's done in either MATLAB or COMSOL. And then that -- the beamforming weights that are done based on simulations are embedded in the C+ -- with the C++ code, and they are compiled.

09:48:27 And just to make sure that I'm not misleading you, source code itself that people are -- that we're looking at here which is human-readable, you can read it, that's not what's on the chip. What's on the chip is a compiled code. It transforms this into 0s and 1s, something that the machine language can understand.

09:48:50 And so the embedded -- the trans -- the weight -- the weights that are done by the MATLAB simulation, as well as the rest of the source code, is all compiled together and then programmed or loaded on the digital signal processor. And it -- it's that that's executed. It's the firmware on the processor that's executed.

09:49:05 So the beamforming weights are constructed based upon one or more configurations of the microphones. And it includes the angle, the distance, which are applied to the microphone array itself. The angle and the distance of

each mic is displaced from a center.

And it also takes into account the azimuth, the target signal. What angle is the target signal coming in from? And from those beamforming weights, they -- they, as well as the rest of the source code, is then compiled, put into for execution by the digital signal processor, and then when the Echo product receives the wake word -- receives the target signal, it will then process that signal with the other signals that are being filtered out, such as ambient noise or disparate signals.

And from that, down to the Fast Fourier Transform, using the samples, it will then create what it aligns to be the six beams based upon the weights that have been previously defined. And it creates those six beams, and that's done by -- and that's all done with -- based on determining the delay.

Q. And does this claim limitation require anything further, sir?

A. No, it does not.

MR. RUBINO: If we could have Slide 54, please. Now, Slide 56, Mr. -- 55, Mr. Thompson. Thank you.

Q. (By Mr. Rubino) Mr. McAlexander, what's the next limitation of this claim?

A. The next limitation which I've bracketed as No. [D] -- as letter [D]. It's: Estimating a spatial location of

09:50:53 said target sound signal from said received sound signals
09:50:57 by said sound source localization unit.
09:50:59 Now we're beginning to invoke one of those units
09:51:02 that's located within the digital signal processor, which
09:51:04 we've identified before, and the local -- the sound source
09:51:10 localization unit that estimates the spatial location of
09:51:13 the target sound signal.
09:51:13 Q. And do you have a high-level depiction of how this is
09:51:19 met, sir?
09:51:20 A. Yes, I do.
09:51:23 MR. RUBINO: If we could please go to PTX-1377 at
09:51:39 Page 62, please.
09:51:41 Q. (By Mr. Rubino) Sir, can you explain this figure
09:51:44 again?
09:51:44 A. Yes. This is in the context of estimating a spatial
09:51:46 location. I had indicated earlier that the voice that's
09:51:49 communicated from the person is going to follow a
09:51:53 particular azimuth, a direction into the microphone array.
09:51:59 And when we look at that, we will see that clearly
09:52:06 the microphone array is -- each one of the microphones is
09:52:09 going to pick up that incoming sound signal. But when it
09:52:14 goes through the process of reaching the determination
09:52:16 based upon the delay, it will determine which microphone is
09:52:21 actually showing the least delay. And in this case, that's
09:52:25 what is shown here as the one labeled "First."

09:52:31 As I had indicated, there is an indication that a number of the Amazon products give by providing a light that lights up when it detects some signal. And then when it determines which signal -- where the signal direction is coming from from the target sound signal, it will then lightly color blue that particular one, which is based on the particular sound source that's coming in -- or the signal that's coming in.

09:53:00 Q. Sir, did you look at any code -- review any code that helped you render your opinions here in this limitation?

09:53:06 A. Yes. Yes, both Doppler and MPAF.

09:53:09 MR. RUBINO: Can we please go to Slide 56 of the Plaintiff's demonstratives? Thank you, Mr. Thompson.

09:53:13 Q. (By Mr. Rubino) Mr. McAlexander, what did you describe in this demonstrative?

09:53:28 A. What I described in this demonstrative is some of the C++ code modules that I reviewed that I believed were strongly applicable to this particular claim element. And I identified the /BeamSelector/BeamSelectorController.cpp.

09:53:36 And several others are identified here, EnergyBeamSelector, ReferenceBeamSelecor, ReferenceTargetBeamSelector, SIRBeamSelector, SNRBeamSelector, the UlpwwwBeamSelector. But all of these are in the module code that's directed to beam selector.

09:54:05 So this is selecting the beam of the six that are

generated, for instance, for a seven-microphone array.
Q. And you said beam selector, what -- what does bream selector mean?
A. That means that you will generate upon speech for a given microphone array -- by example, given seven microphones will generate six beams. So when the spoken word is heard, there will be six beams that are generated.
In the code for the beam selector, those seven input beams will be -- there will be a decision process made as to which one of those is oriented in the -- in the closest direction to the incoming signal. And that -- that one beam will be selected out of the seven. So however many beams are being identified, only one will come out of the BeamSelector code.
MR. RUBINO: Can we please have PTX-359, Mr. Thompson?
Q. (By Mr. Rubino) Mr. McAlexander, do you recognize this document?
A. Yes, I do.
Q. What is this document in the context of that beam selector?
A. It's called the SNR-based beam selector. SNR is signal-to-noise ratio. And part of the process that the Amazon devices goes through is when it -- when a determine -- in a determination step of which beam to

09:55:26 select, one of the algorithms that it -- that it explores
09:55:31 and uses is called the algorithm that's based on
09:55:34 signal-to-noise ratio.
09:55:36 You can look at that as there's a lot of noise
09:55:39 that goes on in an environment. And that creates what we
09:55:42 call a noise flow. And what you're looking at is the
09:55:42 signal that gets out of that noise flow that sits above it.
09:55:48 And so by evaluating the power gradients and the
09:55:51 delay -- the time of delay of the signal arrival, you can
09:55:55 select which beam is the most elevated out of that noise
09:56:00 flow, which one can I really discriminate the best. And as
09:56:04 a result, you select that based upon signal-to-noise ratio.
09:56:07 And the block diagram shows that when you have --
09:56:12 you evaluate the beam energies and you find the one that's
09:56:16 got the best signal-to-noise ratio, one that has the
09:56:21 highest signal-to-noise ratio, then that's the beam that
09:56:26 you select.
09:56:26 This last block called hangover -- excuse me.
09:56:36 Built into the code is a parameter called hangover. And
09:56:40 this says that once the beam is selected, you will stay
09:56:42 with that beam for a certain period of time. That period
09:56:44 of time may be, as I recall, 120 milliseconds.
09:56:49 So it just -- it keeps you from -- literally just
09:56:51 oscillating from beam to beam to beam. It selects a beam
09:56:56 and holds on to it or hangs on to that for some period of

09:57:01 time. And then if it reevaluates the sound that's coming
09:57:03 in and finds that the -- that the person has moved, then it
09:57:06 may select another beam that's more closely associated with
09:57:09 that particular direction.
09:57:16 Q. Sir, do you have any other types of documents that you
09:57:20 looked at for this particular limitation?
09:57:21 A. Yes, I did.
09:57:22 Q. Any publications that you reviewed?
09:57:24 A. There is -- there are certain publications, but one in
09:57:27 particular I wanted to bring to the jury's attention today.
09:57:30 MR. RUBINO: Can we please have PTX-301?
09:57:33 Q. (By Mr. Rubino) What is this publication, sir?
09:57:35 A. This publication is entitled: Multichannel Audio
09:57:45 Front-End for Far-Field Automatic Speech Recognition. And
09:57:52 several of the inventors are listed at the top, including
09:57:56 Chhetri and Hilmes.
09:57:57 Q. And so Mr. Chhetri and Mr. Hilmes are authors of this
09:58:00 publication; is that what you said?
09:58:01 A. That's correct, and it's authored by them as Amazon
09:58:04 employees. It's Amazon Incorporated.
09:58:07 Q. And if you look at the third line, can you please
09:58:10 read -- well, can you please read the first three lines of
09:58:12 the abstract into the record?
09:58:13 A. Certainly.
09:58:14 It says: Far-field speech recognition -- and just

to make sure we're clear, I mentioned this Friday, but when you're speaking into your smartphone, that's near-field, and so the parameters there are very different. But when it's far-field and you've got the potential of extraneous information, it requires a whole -- an entirely different set of algorithms to -- to be able to detect what you're looking for.

So: Far-field automatic speech recognition -- ASR is the acronym for automatic speech recognition -- is a key enabling technology that allows untethered and natural voice interaction between users and Amazon Echo family of products.

MR. RUBINO: Can we zoom out, please, Mr. Thompson?

Q. (By Mr. Rubino) And so what is your understanding about the types of products this article is talking about?

A. It's talking about the Amazon products.

Q. And what's the date in this article, sir?

A. The date is listed as 2018, and it's a conference paper that was presented at the 26th European Signal Processing Conference.

MR. RUBINO: Can we please go to Section 4.3 of the document, Mr. Thompson? That's on Page ending in 55 -- fifth page of the document -- fourth page of the document. Thank you, sir.

Q. (By Mr. Rubino) Mr. McAlexander, what does this section of the document show you?
A. This section is Section 4.3 is identified as Sound Source Localization, and this sometimes is referred to as SSL.
And you'll notice in the first line it talks about look-direction. The user's look-direction identifies that: Knowledge of the user's look-direction is very important for effective beamforming; for Echo products, we need to estimate the look-direction from the microphone array signals. And one of the well-known and robust SSL algorithms is the steered response power. SRP is the acronym for that.
So what this indicates in this paper, it confirms and it -- it actually corroborates other information I've seen, as well as the source code, that the Echo products do use source signal -- excuse me, sound source localization, and this is specifically for the purpose of estimating the spatial location of the target sound signal that's coming in.
Q. So does this limitation for the claim require anything else, sir?
A. No, it does not.
MR. RUBINO: If we could please go back to Slide 57, Mr. Thompson. 58, please, Mr. Thompson. Thank

10:01:14 you.
10:01:14 Q. (By Mr. Rubino) How about the next limitation of the
10:01:16 claim, Mr. McAlexander?
10:01:18 A. The next limitation of the claim reads: Performing
10:01:26 adaptive beamforming for steering a directivity pattern of
10:01:29 said array of said sound sensors in a direction of said
10:01:33 spatial location of said target sound signal by said
10:01:36 adaptive beamforming unit, wherein said adaptive
10:01:38 beamforming unit enhances said target sound signal and
10:01:41 partially suppresses said ambient noise signals.
10:01:46 Q. And is there a construction of a term here, sir?
10:01:49 A. There is plain and ordinary meaning construction for
10:01:52 "steering a directivity pattern."
10:01:57 MR. RUBINO: Can we also look back at Slide 14,
10:02:01 please, Mr. Thompson?
10:02:07 Q. (By Mr. Rubino) Is there a construction for "adaptive
10:02:09 beamforming" that you applied, sir?
10:02:10 A. There is a construction that is identified as: A
10:02:13 beamforming process where the directivity pattern of the
10:02:17 microphone array is capable of being adaptively steered in
10:02:21 the direction of a target sound signal emitted by a target
10:02:24 sound source in motion.
10:02:27 Q. Did you find that in the Amazon products, sir?
10:02:30 A. Yes, I did.
10:02:36 MR. RUBINO: Could you please put up, Mr. --

10:02:39 Slide PTX-12, please, Mr. Thompson? Plaintiff's Exhibit
10:02:45 12. Thank you. And can we turn to Slide -- Page 2 of the
10:02:49 document?
10:02:52 Q. (By Mr. Rubino) Mr. McAlexander, how is this
10:02:54 limitation met?
10:03:01 A. This is for the Audio Front End Architecture for the
10:03:09 Sonar. And it states in the first paragraph about the fact
10:03:16 that the algorithms perform critical functions required for
10:03:21 far-field automatic speech recognition, ASR, and
10:03:25 high-quality audio feedback. And the diagrams that are
10:03:32 shown on the pages that follow show how that operates.
10:03:35 Q. Is this the diagram you were speaking of, sir?
10:03:38 A. This is the diagram for Doppler. And the -- the
10:03:44 particular parameter we are identifying are -- is -- is the
10:03:55 estimation of the spatial location. And that particular
10:03:59 done -- is done through the Main Beamformer Selector, and
10:04:08 that is where the selection takes place.
10:04:10 Q. How about for the MPAF class of products, sir?
10:04:13 A. Similarly.
10:04:14 MR. RUBINO: Can we have Plaintiff's 79, please?
10:04:21 Q. (By Mr. Rubino) Sir, what is this document?
10:04:23 A. This document is for the Biscuit MPAF. It's
10:04:29 Plaintiff's Exhibit 79. Again, identifying the
10:04:32 architecture for the audio front end.
10:04:33 MR. RUBINO: Can we have Page 5 of the document,

10:04:36 please, Mr. Thompson? If you can zoom in on the very top line.
10:04:41 Q. (By Mr. Rubino) Mr. McAlexander, what is this depicting?
10:04:44 A. This is -- detects -- is depicting for the MPAF similar to what we saw in Doppler. There is a beam selection that is occurring. It more informatively shows us that the adaptive beamformer is used as a part of that process.
10:05:02 Q. Now, that says ABF. ABF stands for adaptive beamformer?
10:05:11 A. Adaptive beamforming, yes.
10:05:14 Q. And is that in the document somewhere?
10:05:18 A. Yes.
10:05:18 MR. RUBINO: If we could go to the second page of the document, please, Mr. Thompson. Very top two lines. Thank you.
10:05:23 Q. (By Mr. Rubino) Mr. McAlexander, what does this show in the document?
10:05:27 A. Well, it shows two things. One is the acronym ABF does stand for adaptive beamforming, and it includes fixed beamforming. So there's a fixed beamforming component that's in the formation of the adaptive beamforming.
10:05:42 Q. And so this adaptive beamforming unit in the accused products, what exactly does it do to the beam?
10:05:47 A. What it does is it -- if I can go back, the fixed

10:05:52 beamforming identifies the beams that are selectable. And the -- the selector selects the one from that group. Then that points the beam in the direction of the sound signal that's coming in.

10:06:07 Now, if the -- if the target themselves, if the person moves to a different location in the room and is still speaking, then every 120 milliseconds it allows in the code for that beam to be steered and follow the motion of that person.

10:06:23 And so, here, the adaptive beamforming is, according to the definition, it's for steering the pattern, and that's done if the target -- if the sound source is moving.

10:06:38 And in this case, that's exactly what the adaptive beamforming does. It will actually follow the -- the user of the system that's speaking. And that is shown as at least an indicator. For instance, if you look at the light -- the light blue that comes on the LED ring that points in the direction of the user, if the user moves, that light will follow.

10:07:00 Q. Thank you, sir.

10:07:01 And you said something about code. Did you find this in the code as well?

10:07:06 A. Oh, yes.

10:07:07 MR. RUBINO: Can we have Slide 61 of Plaintiff's,

10:07:16 please?
10:07:16 Q. (By Mr. Rubino) Is this some of the code you looked at
10:07:18 for this limitation, sir?
10:07:20 A. Yes, in this -- this particular code for the MPAF is
10:07:23 located in the AdaptiveBeamFormer module. It
10:07:33 specifically -- it identifies specific code that's within
10:07:35 those modules.
10:07:35 Q. So does this limitation require anything further, sir?
10:07:40 A. No, it does not.
10:07:42 MR. RUBINO: Can we please go to Slide 63?
10:07:44 Q. (By Mr. Rubino) How about the final limitation of
10:07:46 Claim 1, Mr. McAlexander?
10:07:49 A. The final limitation reads: Suppressing said ambient
10:07:55 noise signals by said noise reduction unit for further
10:07:58 enhancing said target sound signal.
10:08:05 Q. And how is this met by the accused products, sir?
10:08:09 A. The accused products, in addition to having the
10:08:19 beamforming algorithms and the adaptive beamforming
10:08:21 algorithms and the sound source localization algorithms, in
10:08:27 addition to having those, they also have a number of units
10:08:30 that address noise, some of which I mentioned earlier like
10:08:35 the high pass filter that eliminates everything below 80
10:08:43 hertz. It also has others called echo cancellation.
10:08:46 MR. RUBINO: If we can please have Plaintiff's 12,
10:08:51 again, Mr. Thompson, and Figure 1.

10:08:58 Q. (By Mr. Rubino) Mr. McAlexander, can you explain where that voice cancellation is in Figure 1?
10:09:05 A. Well, it's in a number of places. There is high pass filtering that is performed, as I've shown previously. And also the modification of the beam is highlighted with -- it's called an Acoustic Echo Canceler. So if there is echo that is determined to be present, it will cancel that.
10:09:29 Sometimes when something is spoken, there may be some artifacts that come through the speaker system, and that would be an echo of what is actually spoken. And so the echo canceler will go in and determine that that is a residual, and cancel that information out.
10:09:51 Q. Now, does Amazon itself say anything about how noise cancellation is done?
10:09:55 A. Yes.
10:09:56 MR. RUBINO: Can we please have Slide 64?
10:09:59 Q. (By Mr. Rubino) What does this demonstrative show, sir?
10:10:07 A. This demonstrative is Plaintiff's Exhibit 399. It states: Introducing Sonar, Higher Performance Follow-up to Amazon's Groundbreaking Voice-Controlled, Cloud Connected Home System. And it states that Sonar boosts -- or boasts, excuse me, improved speech recognition, enhanced music playback and expanded voice control for the home.
10:10:35 Q. Do you see anything about noise reduction in the second

paragraph, sir?
A. Yes. When you look at the -- at the second paragraph under Sonar, it provides the best far-field speech recognition. It states that: Sonar engages a series of advanced audio algorithm to drastically improve the far-field speech recognition allowing Alexa to process customer speech with greater accuracy than ever before.
It goes on to say: Now Alexa is able to discern the customer's voice more clearly even while a customer is rocking out to Lady Gaga, hosting a loud dinner party, or watching a movie on a 5.1 surround home stereo -- surround sound home theater system.
So it goes on to say that: It's very responsive when spoken to. It's better able to tune out even common household noises, such as dishwashers and fans.
So this goes to the cancellation and the noise abatement algorithms that are built into the accused device so that it meets what is required for suppressing the ambient noise signals.
Q. And so does this limitation require anything further, sir?
A. No, it does not.
MR. RUBINO: Can we have Slide 66, please?
Q. (By Mr. Rubino) So does the claim as a whole require anything further, sir, from the Amazon products?

10:12:01 A. No, it does not. I have gone through each one of the claim elements from the preamble and each one of the steps, and I've identified evidence, based on the information provided, in terms of code, product, teardown of product, deposition testimony, documents that have been produced in this matter, technical documents, specifications on digital signal processors, and I've identified all of that. And my opinion is that each one of the accused products practices each one of the elements of Claim 1 of the '049 patent.

10:12:42 Q. What about Claim 8, do you also have an opinion about Claim 8, sir?

10:12:46 A. Yes, I do.

10:12:46 THE COURT: Let me interrupt right here. Before we get into Claim 8, we're going to use this opportunity to take a recess.

10:12:53 Ladies and gentlemen of the jury, if you'll simply close your notebooks and leave them in your chairs, follow all the instructions I've given you, including not to discuss the case among yourselves, and we'll be back shortly to continue with the next asserted claim through this witness.

10:13:09 The jury is excused for recess at this time.

10:13:11 COURT SECURITY OFFICER: All rise.

10:13:13 (Jury out.)

10:13:37 THE COURT: Be seated, please.

10:13:38 Counsel, just for your information, according to the Court's records, Plaintiff has 6 hours and 17 minutes remaining of their designated trial time. And Defendant has 9 hours and 3 minutes remaining from their trial time.

10:13:58 Also, over the weekend, the Court reviewed the joint submission from the parties of the proposed final jury instructions and verdict form.

10:14:12 I'm persuaded that it would be appropriate and beneficial to the Court for the parties to meet and confer, and submit a reviewed and updated proposed final jury instruction and verdict form.

10:14:27 Consequently, I'm going to direct that you jointly do that and submit it in Word format to the Court for further review not later than 3:00 p.m. tomorrow.

10:14:37 All right. We will take approximately a 10 or 12-minute recess, and then we'll return and continue with this witness.

10:14:46 The Court stands in recess.

10:14:47 COURT SECURITY OFFICER: All rise.

10:14:48 (Recess.)

10:14:49 (Jury out.)

10:14:49 COURT SECURITY OFFICER: All rise.

10:34:01 THE COURT: Be seated, please.

10:34:01 Mr. Rubino, are you prepared to continue with your direct examination?

MR. RUBINO: Yes, Your Honor.

THE COURT: You may return to the podium.

Mr. Johnston, if you'd bring in the jury, please.

COURT SECURITY OFFICER: All rise.

(Jury in.)

THE COURT: Please be seated.

Counsel, you may continue with your direct examination of the witness.

MR. RUBINO: Thank you, Your Honor.

Mr. Thompson, could we please have Slide 67?

Q. (By Mr. Rubino) Mr. McAlexander, did you form an opinion with regard to Claim 8 of the '049 patent?

A. Yes, I did.

Q. What is your opinion about this claim?

A. My opinion is that the products that use the MPAF code practice Claim 8.

Q. Can you read Claim 8 into the record, please?

A. Yes. And if I may have the graphics -- the last four words on the first line please, put a red line through those. They're repetitious.

So the Claim 8 reads: The method of Claim 1, wherein said noise reduction unit performs noise reduction in a plurality of frequency subbands, wherein said frequency subbands are employed by an analysis filter bank of said adaptive beamforming unit for subband adaptive

beamforming.

MR. RUBINO: Can we please go to Slide 68? Thank you, Mr. Thompson.

Q. (By Mr. Rubino) Mr. McAlexander, how is this claim infringed?

A. What I'm showing here is, when I looked in the modules of code specifically for the MPAF, Multi Platform Audio Framework, there are sets of code that are loaded in a module called Subband.

And each one of these subband modules of code include FilterBanks, and the FilterBank -- FilterBanks reply on a subband basis. So this is taking the band frequencies, dividing it, and they're performing the Fast Fourier Transforms in each one of those bands in parallel.

Q. And, sir, does this limitation require anything further?

A. No, it does not.

MR. RUBINO: If we could go to Slide 60, please -- Slide 60, please, Mr. Thompson.

Q. (By Mr. Rubino) So, Mr. McAlexander, can you summarize your opinions here with regard to infringement?

A. Yes. I have covered each one of these accused products and have identified evidence that I believe supports a finding of infringement for each one of them.

Specifically, my finding is any end user that uses

10:37:29 the product, just by the mere speaking of the wake word,
10:37:33 infringes Claim 1. And for the MPAF, it infringes Claims 1
10:37:39 and 8.
10:37:41 And, secondly, Amazon, by offering to sell and
10:37:47 selling the product, indirectly infringes each one of these
10:37:51 claims, as well.
10:37:54 Q. Do you have any additional evidence of this indirect
10:37:58 infringement, sir?
10:37:58 A. Yes.
10:38:00 MR. RUBINO: If we could please go to Slide 70.
10:38:07 Q. (By Mr. Rubino) Mr. McAlexander, can you explain
10:38:10 what's on this slide?
10:38:11 A. We have seen this one before. This is Plaintiff's
10:38:14 Exhibit 1372, and this is Amazon's information to the users
10:38:23 on how to set up their Echo. And, specifically, they
10:38:27 indicate the steps that one goes through to set it up, and
10:38:31 then follow the instructions.
10:38:36 Q. Mr. McAlexander, did you perform any analysis of the
10:38:41 value of the patented technology?
10:38:42 A. Yes, I did.
10:38:45 Q. And what is that analysis?
10:38:46 A. Well, with regard to the -- the value of the patented
10:38:52 technology, I first looked at the claim itself, you know,
10:38:57 what is the claim directed to? And the claim is directed
10:39:00 to a system that follows certain steps and protocols. And

those steps are how to discriminate, how to detect and pull out of an environmental sound the particular target sound signal.

So if a spoken word or words are said, the question is, how does one go about doing that to -- to specifically identify the word that you're looking for and pull it out of the rest of the environment? That's what this claim is directed to.

So when I look at the way -- the steps by which it enacts that, it specifically is addressing key factors that are necessary to discriminate and determine that particular sound signal, some of which are providing filtration so that you have the ability to fil -- filter out disparate signals, filter out ambient noise so that you can more clearly discriminate where the signal is coming from.

Secondly, not only does the patented invention, in terms of the value it adds for Amazon or others that use it, the ability to discriminate, but it also aligns the beam -- the beampattern in such a way that it promotes the -- the beam and the direction in which the sound is coming from. So it does that with the beamforming algorithms.

But the key thing that I find also is an attribute is it does it in a far-field environment, that it can discriminate between sounds that are reflecting off of

walls. You know, when I'm saying something, the words that I'm saying comes from me, but the microphone is projecting it through the speakers.

And so if -- if one is to align a beam with a direction that the sound is coming from, it must somehow be able to get rid of the sound that's coming from the speakers and direct it just to me.

And to do that, that requires some high-level math in order to create that discrimination and eliminate the other reflections that are occurring.

Also, if music is playing in the background, the -- the Amazon systems are able to discriminate the spoken word when it is -- the -- the wake word. It can discriminate that, even with music and other background noise playing.

So great value added, I believe, because the patented invention would embody -- as embodied in the Amazon products, enables it to, with clarity, discriminate the spoken wake word from everything else that's happening. And that's regardless of whether it's in a room with reflecting walls or even far-field -- when it's outside when there's no reflection but you have other ambient disturbances going on.

Q. Thank you, sir.

And how does that compare to the other features of

10:41:59 that product?

10:42:00 A. Well, I think that as it -- as that compares to other features of the product, I would -- I would characterize it as it -- it provides what I would call a fundamental value that advances this product. And I would put it on par -- it's -- it's at least as important as the rest of the -- the features put together.

10:42:25 So much of the rest of the features depend upon voice discrimination and being able to select the words that are said. Even the wake word, which is the only thing required when spoken for the invention to be practiced -- for Claim 1 and Claim 8 to be practiced, the only thing that's required, is that the ability to discriminate that word also allows the -- the front end to be able to discriminate other words that could be used in command structure and everything.

10:42:54 So the weight of what is done to work on the wake -- wake word applies to the other parts of the features, as well.

10:43:01 So I put this on par with -- it's -- it's a fundamental value that's added to the advancement of the product. It's -- it's at least as important as the other technology that the product has.

10:43:11 Q. And, sir, what about any non-infringing uses of the product?

10:43:18 A. Well, non-infringing -- if you're asking the question about other ways in which it could be done, my understanding of non-infringing uses is that there's two -- two parts to this.

10:43:33 One is, you have to be able -- if it's a non-infringing alternative, it has to be able to achieve the same result.

10:43:41 And so there are other alternatives that you can do by touching a screen or something of that nature and invoking commands, like pushing buttons. But that requires you to be up close to the product. You can't move around.

10:43:58 So the result is not the same. You've got to be able to be free in the environment and have that word understood.

10:44:03 In terms of non-infringing uses of the product as sold, I don't see any, because it's all based upon as soon as the wake word is spoken, which is what starts the -- the Amazon products to operate and -- and to respond, the invention is practiced.

10:44:20 Q. What about alternatives to the product?

10:44:23 A. Again, the only alternatives that I see is you can do things by defeating some of the -- the filtration or not having the beams steer in the direction of the person moving. And you can do it by -- by instead of doing samples, you could do it with time. You could do a lot of

10:44:45 other things.

10:44:46 But the bottom line is you don't achieve the result that this achieves. And if you don't achieve the result, then that's not an alternative design.

10:44:54 Q. So what was your conclusion about these technical portions of this product, sir?

10:45:01 A. My conclusion, as far as -- the technical portions of the products; is that what you said?

10:45:08 Q. Yes, sir.

10:45:09 A. Well, my evaluation, certainly from a technical standpoint -- first of all, it's -- each one of these systems is a very smart system. Yeah, there's a lot of other things that it does, but bottom line is built into the -- each one of these containers is a very smart system.

10:45:28 Each one of them has a very sophisticated microphone array, very sophisticated code -- that's firmware that -- that it executes -- very sophisticated processors. Using the ARM core Neon-type processors is a very sophisticated processor. And each one of these is capable of doing high-level mathematical operations. It's specifically tailored for the -- what it does.

10:45:57 I think it also gives great freedom to the user to know that -- whether you are near it or far away from it it can still pick up and discern, is great value added from a technology standpoint, because there's a lot of things you

10:46:11 can do with it once it can discern the words that you're
10:46:15 saying.
10:46:15 Q. Thank you, Mr. McAlexander.
10:46:22 MR. RUBINO: I pass the witness.
10:46:23 THE COURT: Cross-examination by the Defendant?
10:46:25 MR. HADDEN: Yes, Your Honor.
10:46:26 May we approach with some binders, Your Honor?
10:46:29 THE COURT: You have leave to distribute binders.
10:46:31 MR. HADDEN: Thank you, Your Honor.
10:46:52 THE COURT: All right. Counsel, you may proceed
10:46:54 with cross-examination.
10:46:55 MR. HADDEN: Thank you, Your Honor.
10:46:55 CROSS-EXAMINATION
10:46:57 BY MR. HADDEN:
10:46:57 Q. Good morning, Mr. McAlexander.
10:46:58 A. Good morning, sir.
10:46:59 Q. You're not an expert in adaptive beamforming, are you?
10:47:02 A. Not in adaptive beamforming, per se, no.
10:47:05 Q. And you've not authored any publications regarding
10:47:09 adaptive beamforming, have you?
10:47:10 A. That's correct, no.
10:47:11 Q. And you've not taught any classes regarding adaptive
10:47:14 beamforming, have you?
10:47:15 A. That's also correct.
10:47:16 Q. In fact, you haven't even taken any classes related to

10:47:20 adaptive beamforming, have you, Mr. McAlexander?
10:47:22 A. Not any official classes, no.
10:47:24 Q. Okay. And, in fact, you have no experience with adaptive beamforming as it is used in these patents, do you, Mr. McAlexander?
10:47:31 A. Not correct. As I mentioned, I have used -- in acoustic design of systems, I have used equipment. I haven't designed the actual construct of the -- the adaptive beamformer itself, but I've used it.
10:47:44 Q. Well, do you recall being -- having your deposition taken in this case, Mr. McAlexander?
10:47:49 A. Certainly.
10:47:49 Q. And do you recall being deposed on June 9, being asked the question: Do you have any experience with adaptive beamforming?
10:48:00 A. Yes. And I believe I answered no.
10:48:03 Q. Excuse me?
10:48:03 A. Yes, I remember that.
10:48:04 Q. And you recall that your answer was: Other than what I've discussed with you in the radar system, the answer is no?
10:48:11 A. Correct.
10:48:12 Q. Not as used in this patent?
10:48:15 No?
10:48:15 A. Correct.

10:48:16 Q. And you're not an expert in sound source localization
10:48:20 either, are you, Mr. McAlexander?
10:48:22 A. No, I expressed in my deposition how I have done this
10:48:26 in terms of system implementation using external
10:48:30 microphones, but not in terms of what we're talking about
10:48:32 here, no.
10:48:33 Q. So you've not authored any publications related to
10:48:38 sound source localization, have you?
10:48:39 A. No, I have not.
10:48:40 Q. And you've not taught any classes related to sound
10:48:45 source localization either, have you?
10:48:45 A. That's correct, I have not.
10:48:46 Q. And you've not taken any courses related to sound
10:48:50 source localization, have you, Mr. McAlexander?
10:48:51 A. That is correct, I have not.
10:48:53 Q. Okay. And you, in fact, have no experience with sound
10:48:59 source localization as defined in and used in this patent,
10:49:05 do you, Mr. McAlexander?
10:49:05 A. That is correct.
10:49:06 Q. And you're not an expert on microphone arrays, either,
10:49:09 are you, Mr. McAlexander?
10:49:10 A. I had expressed in my deposition the fact that I have
10:49:13 done microphone arrays on an analog microphone, but in
10:49:16 terms of this particular aspect, I've not done this myself,
10:49:19 no.

10:49:20 Q. Okay. And you have not authored any publications or technical papers relating to microphone arrays, have you?
10:49:25 A. No, I have not.
10:49:26 Q. And you have not taught any courses related to microphone arrays, have you?
10:49:30 A. Correct, I have not.
10:49:32 Q. In fact, you haven't taken any courses related to microphone arrays, have you?
10:49:36 A. Correct, I have not.
10:49:37 Q. Okay. Is it fair to say that you have never worked with, installed, or used a system that has a fixed beamformer, a blocking matrix, and an adaptive filter?
10:49:56 A. I have used the Amazon systems, as I have indicated, so in that case, I have.
10:50:02 Q. Okay. But before you were hired by Mr. Fabricant in this case, is it fair -- fair to say that you had never worked with, installed, or used a system that had a fixed beamformer, a blocking matrix, and an adaptive filter?
10:50:17 A. I've worked with Amazon products before I was hired by Mr. Fabricant.
10:50:21 Q. Okay. Do you recall, again, in your deposition on June 9th being asked the question: Okay. So is it fair to say, though, that you have never worked with, installed, or used a system that has a fixed beamformer, a blocking matrix, and an adaptive filter?

10:50:41 And do you recall answering: I would say the answer to that is yes. I have not.
10:50:49 A. That is correct, other than the Amazon products, that is correct.
10:50:53 MR. HADDEN: Can we pull up the patent, please, Mr. Berk? And can we go to Figure 1, please? Blow that up a little bit, thank you very much.
10:51:12 Q. (By Mr. Hadden) Do you recognize this as Figure 1 from the '049 patent, Mr. McAlexander?
10:51:15 A. Yes, that's correct.
10:51:16 Q. And this is also the figure that's on the front of the patent; isn't that right?
10:51:20 A. That is also correct.
10:51:22 Q. And if we look at this flowchart, and start with Element 102, it says: Receive multiple sound signals from multiple disparate sources by the sound sensors. Do you see that, Mr. McAlexander?
10:51:43 A. Yes, I do.
10:51:43 Q. And then following Step 102 is Step 103, it says: Estimate a spatial location of the target sound signal from the received sound signals by the sound source localization unit.
10:52:00 Do you see that, Mr. McAlexander?
10:52:01 A. Yes, I see that, as well.
10:52:03 Q. Is this describing it after the system of microphones

receives the sounds, it then locates the target sound source using this spatial location -- using this sound source localization unit?
A. Well, this describes Steps 102 and 103. And in this -- in this particular illustration, it -- it shows a method for enhancing the target sound signal performed in these two steps.
Q. And then the step that comes after that in this Figure 104 it describes: Perform adaptive beamforming for steering a directivity pattern of the array of sound sensors in a direction of the spatial location of the target sound signal by the adaptive beamforming unit.
Do you see that, Mr. McAlexander?
A. Yes, I do.
Q. So in this Figure 1, it describes that the target sound source is located, and then a beam is formed in that direction. Isn't that right, Mr. McAlexander?
A. Well, these are -- now you're showing three steps, 102, 103, and 104, and these three steps are performed as a part of this method that is illustrated in Figure 1.
Q. And in this method that is illustrated in Figure 1, the system adaptively steers the beam in the direction of the target sound source that was located using the sound source localization unit. Isn't that correct?
A. Well, the -- the statement that's shown here is three

10:53:43 separate steps, and the way it is diagrammed in this
10:53:47 method, the adaptive beamforming is done in Step 104. Now,
10:53:53 certainly the way it's diagrammed, 103 is -- is shown above
10:53:57 that in this illustration.
10:53:58 Q. Right. It's also shown that there is an error going
10:54:02 from Step 103 to 104. Isn't it typical in a flowchart,
10:54:07 Mr. McAlexander, that when you have an arrow going from one
10:54:10 box to the next, that shows a sequence of steps?
10:54:15 A. It shows a -- it shows steps, and they show that they
10:54:19 are interrelated. It's a typical block diagram formation.
10:54:22 Q. And is it your understanding when you're reading
10:54:24 Figure 1 that Step 104 followed Step 103?
10:54:29 A. Well, you can't write two words on top of each other,
10:54:33 so, yes, one -- one follows the other so that you can
10:54:36 discern 103 from 104.
10:54:39 Q. And you understood from reading this patent that it
10:54:42 describes a system in which the target sound source is
10:54:47 located and then a beam is formed to point in the direction
10:54:50 of that source; isn't that right?
10:54:51 A. That's my understanding. That's one of the
10:54:54 descriptions given in the spec, yes.
10:54:57 MR. HADDEN: And if we look at Figure 16,
10:55:00 Mr. Berk. So if we can blow this up.
10:55:07 Q. (By Mr. Hadden) Figure 16E and 16F show the result of
10:55:11 that process where a beam is being formed to point in a

10:55:16 particular direction. Isn't that right, Mr. McAlexander?
10:55:19 A. Figure 16E through 16L, the specification states, is
10:55:32 that it is an exemplary illustrate -- illustrated graphic
10:55:37 representation showing the directivity patterns of the
10:55:41 eight-sensor microphone array in Figure 16A. So this is an
10:55:46 exemplary illustration of that particular configuration.
10:55:49 Q. And those -- those patterns that are --
10:55:53 MR. HADDEN: Maybe we could blow up one of the
10:55:56 beampatterns in 16E, Mr. Berk.
10:56:03 Q. (By Mr. Hadden) So that kind of funny-shaped blob in
10:56:07 the middle of the circle is the beam that is being formed.
10:56:11 And in this case it's being formed to point at a target
10:56:15 source. It would be approximately 15 degrees in this
10:56:18 diagram. Isn't that right, Mr. McAlexander?
10:56:20 A. Yes. The diagram that showed an illustration here in
10:56:24 this particular microphone beampattern, there are a number
10:56:28 of different beams. And this is basically identifying one
10:56:34 that's pointed in the direction of the target sound source.
10:56:36 Q. Okay. And to form the beams like we see here in 16.1
10:56:47 that are pointed at a particular target sound source, the
10:56:54 patent first determines a series of delays; isn't that
10:57:00 correct, Mr. McAlexander?
10:57:01 A. When you say the patent, are you talking about the
10:57:04 claims?
10:57:04 Q. Both. But let's start with the written description in

the specification.
A. Some of the illustrations that -- that are defined is that -- in terms of the beamforming algorithm, you will define the beampatterns, and then from that you will select the one that is pointed in the direction of the target sound source. That's -- that's the illustration that's given in the patent.
Q. Well, in fact, the way the patent works is you determine delays to the different microphones based on the location of the target sound source, and from those delays, you calculate the beampattern. Isn't that correct, Mr. McAlexander?
A. Can you say that again, please?
Q. Sure. You form this beampattern, like we see in Figure 16E, the patent as it describes, first determines that the target sound source is at 15 degrees. And then using that information, it calculates the delays to each of the microphones from the origin. And then it uses those delays to form those weights you talked about that actually caused this beam to have this shape. Isn't that correct?
A. I think your general description of the specification illustration is correct.
MR. HADDEN: Can we go to Figure 5, please, Mr. Berk? So if we blow up Figure 5.
Q. (By Mr. Hadden) This is an illustration showing how

10:58:37 those delays are calculated, isn't it, Mr. McAlexander?
10:58:40 A. Well, it shows how those delays are determined, yes.
10:58:47 And it's -- it's an example, and it actually depends on
10:58:54 Figure 4 of this illustration, or the delay-and-sum in this
10:58:59 particular Figure 4 is the technique that is used.
10:59:01 Q. So just to put a -- on Figure 5, that arrow that says
10:59:05 target sound signal, that is pointing to the target that is
10:59:11 emitting the sound that we are trying to capture; isn't
10:59:11 that correct?
10:59:21 A. I believe the target sound signal arrow is showing the
10:59:25 direction -- the azimuth from which the target sound signal
10:59:27 is coming. So that gen -- that is where it's coming from,
10:59:30 and, obviously, it's continuing to pass through there, so
10:59:34 it's going in that direction, too.
10:59:35 Q. Sure. So in your example of someone speaking the wake
10:59:40 word "Alexa," that arrow would be pointing to that person,
10:59:44 right?
10:59:44 A. Again, the arrow -- the arrow -- the beamformer, if --
10:59:56 if that's the target sound signal direction, if it's coming
10:59:58 from the top right in that quadrant, then the beamformer
11:00:03 would be then directed to steer and move in that direction.
11:00:08 Q. So if the person is speaking "Alexa" is the target
11:00:13 sound source in the example, that arrow would be an arrow
11:00:16 pointing to that person; isn't that correct?
11:00:24 A. If -- whatever direction the angle -- whatever

direction that the arrow was coming from, that would be the originating direction from the target sound signal, and it's going to pass through and pass by the microphones and continue on. So there would be associated with that arrow a direction that is associated with the target sound signal.

Q. Right. And the direction to the target sound signal would be, in my example, the person speaking the wake word "Alexa." That is required to calculate or determine these delays, right?

A. As far as determining the incident azimuth, yes, then there would be a beam finder -- beamformer -- beamforming code that would be done for delay-and-sum that would identify that based on the weights of the beams, the direction at which the one that has the primary above the signal-to-noise ratio, and then, yes, from that you can make a selection that would be associated with that beam direction.

MR. HADDEN: Move to strike as non-responsive, Your Honor.

THE COURT: Overruled.

Q. (By Mr. Hadden) Focusing just on the delay calculation, Mr. McAlexander, to calculate the delays as shown in Figure 5 you have to know the angle to the target sound source; isn't that correct?

11:02:05 A. To calculate the delay, you would have to know the angle?
11:02:09 Q. Yes, Mr. McAlexander.
11:02:11 A. No. Your calculation of the delay goes into a delay algorithm that does a filter-and-sum or a delay-and-sum in the Fast Fourier Transform area. And you're evaluating the -- the components that are coming -- that are determined for each of the beams.
11:02:30 And so in that beamforming activity, there may be a particular incident beam that you will eventually select. But in terms of the initial calculation, your calculation is based on the weight and the beams itself. And from that, you will then make a determination from the overall weights and the overall delays that are determined. Then you will select from that the direction of the beam.
11:02:58 Q. Well, let's just focus on Figure 5. Do you see the delays in Figure 5, Mr. McAlexander?
11:03:04 A. Yes.
11:03:08 Q. So those are those tau 1 and tau 3?
11:03:11 A. It's not very easy to read. I hope the jury can see that. But, yes, those are the taus. This is a -- the Greek symbol that's used for delay.
11:03:21 Q. And so the delay between Microphone 3, which is M_3, in the origin is tau 3; isn't that right, Mr. McAlexander?
11:03:33 A. That is correct.

11:03:34 Q. And that tau 3 is shown here as a distance; isn't that right?
11:03:39 A. Yes. And that distance -- if there was a sig -- signal that is incident, for instance, on the M_3, then if it's incident perpendicular to M_3, such as coming in here, then it's the distance -- the delay that's determined to be tau 3 would be consistent with the delay from that vector that's coming in directly tang -- or perpendicular to that microphone.
11:04:10 Q. Right. But the tau 3 and -- and just -- just to be clear, can be represented as a distance, because the speed of sound is constant. So time is proportional to distance; isn't that right?
11:04:24 A. Well, they're -- they're different -- those are different parameters, but there is a proportionality that you can derive from that based upon the architectural organization of the microphone array.
11:04:35 Q. The distance between -- a difference -- the difference between a distance -- sorry.
11:04:39 Converting a distance to a time just depends on the speed of sound, doesn't it, Mr. McAlexander?
11:04:47 THE COURT: Counsel, you're going need to speak up.
11:04:50 MR. HADDEN: Sure, I'm sorry.
11:04:51 Q. (By Mr. Hadden) The conversion of tau 3 from a

distance to a time measurement just depends on the speed of sound, doesn't it?
A. Speed of sound is certainly a parameter that is used in that, yes.
Q. Okay. And going back to this figure, tau 3, the actual distance that is shown in Figure 5, depends on that angle theta, which is the azimuth angle to the target sound signal. Isn't that right?
A. That's correct. As I mentioned before, if the -- if the target sound signal is coming in exactly perpendicular to M_3, then the tau that's associated with that would be correlated directly with the actual distance from Microphone 3 to the center.
If it's off azimuth, then there's going to be some difference that has to be performed in making that determination.
Q. So to determine the delay from each of the microphones to the origin of the microphone array, requires knowing that azimuth angle to the target sound signal; isn't that right?
A. For this illustration, that would be correct.
Q. And those delays that are determined as we see in Figure 5 in the patent, those are what are used to then calculate those beamforming weights; isn't that right, Mr. McAlexander?

11:06:24 A. There is -- there are already known weights to be based
11:06:31 on the architecture of the microphone. Then when you have
11:06:37 a sound signal that is coming in that is off azimuth, then
11:06:40 there will be some weight calculations or determinations
11:06:43 that are done based upon the angle of incidence of that
11:06:48 target sound signal.
11:06:57 MR. HADDEN: Can we look at Column 8, Lines 14
11:06:59 through 17, Mr. Berk? After 301, can you blow that up,
11:07:42 Mr. Berk, or highlight it, for any specific microphone
11:07:45 array configuration? No, at Line 14. I'm sorry, Mr. Berk,
11:08:00 Line 14 to 17.
11:08:18 Q. (By Mr. Hadden) So if we look at the specification
11:08:20 here, Mr. McAlexander, it says that: For any specific
11:08:24 microphone array configuration, the parameter that is
11:08:27 defined to achieve the beamformer coefficients is the value
11:08:32 of the tau_n for each sound sensor 301.
11:08:38 Do you see that?
11:08:38 A. Yes. That's in agreement with what I said about this
11:08:41 particular illustration.
11:08:42 Q. Okay. And the beamformer coefficients, those are those
11:08:46 weights that you talked about that multiply the output of
11:08:50 the microphones in the filter-and-sum algorithm; isn't that
11:08:50 correct?
11:08:55 A. That's correct. There's some portions of the
11:08:57 filter-and-sum algorithm that are already known as

11:08:58 coefficients because you already know the architecture of
11:09:01 the -- of the microphone array itself to the center.
11:09:04 And so then when you apply that in situ with a
11:09:10 particular incoming beam, then you have the azimuth portion
11:09:15 of that, and so now you have the full complement of the
11:09:18 coefficients. I agree with that.
11:09:20 Q. And what it says here is that those delays that are
11:09:23 determined are what are used to achieve those beamformer
11:09:26 coefficients; isn't that right?
11:09:28 A. That's what a delay-and-sum beamformer does.
11:09:32 Q. So you need to have a delay before you can create the
11:09:37 beamformer coefficients in the -- in the filter-and-sum
11:09:42 algorithm you talked about, right?
11:09:43 A. No, that's not correct. It's achieve, not create. You
11:09:49 can't achieve unless you have the incoming beam. You have
11:09:52 the azimuth that is a component in that calculation. So
11:09:56 you have some of it that you already know. But when you
11:09:58 have the incoming beam, that then determines which
11:10:02 coefficient you are going to use.
11:10:04 Q. Just to be clear, what the patent says is, to achieve
11:10:09 beamer coefficients, use the value tau. And that's the
11:10:14 delay, right?
11:10:14 A. That's what it says, but, again, understand that the
11:10:16 beam -- the coefficients in terms of the actual placement
11:10:18 of the microphones is already in there; it's already

11:10:20 defined. So when you have the azimuth, that gives you the -- what I call the in situ. That's the real-time, here it is.

11:10:28 So then you can create the actual coefficients at that point that are associated with that particular beam coming in. That's what it says, and that's what it does.

11:10:37 MR. HADDEN: May I put up a board, Your Honor?

11:10:39 THE COURT: You may.

11:11:02 Can you see this, Mr. McAlexander?

11:11:05 THE WITNESS: For the most part.

11:11:07 THE COURT: Pull that forward, if you will, counsel.

11:11:10 MR. HADDEN: Forward a little bit.

11:11:12 Q. (By Mr. Hadden) Is that better, sir?

11:11:20 A. To the left a little.

11:11:21 THE COURT: I tell you what, why don't you stand up.

11:11:24 THE WITNESS: Can I stand up?

11:11:25 THE COURT: You may stand up if -- when necessary.

11:11:25 THE WITNESS: Thank you.

11:11:25 THE COURT: All right. Let's proceed, counsel.

11:11:30 MR. HADDEN: Thank you, Your Honor.

11:11:30 Q. (By Mr. Hadden) Now, we talked about the specification. When we talked about --

11:11:31 MR. HADDEN: Can we also put up Claim 1 on the

11:11:35 screen, Mr. Berk?
11:11:40 Q. (By Mr. Hadden) And before we focus in on the determining step, you agree, Mr. McAlexander, that each of these steps has to be performed by the device when it's in use; isn't that right?
11:11:50 A. Yes, each of the steps -- each of the steps has to be performed by an accused device.
11:11:57 Q. Okay. And when you say "by an accused device," you mean by an Echo when a customer is using it, correct?
11:12:05 A. Correct. As I've stated on my direct, when the customer is using it, speaks the wake word, all of these steps are performed.
11:12:14 Q. And that includes this determining step; is that correct, Mr. McAlexander?
11:12:19 A. Yes, as I explained on my direct.
11:12:21 Q. So the determining step has to be performed on the device while it's in operation, correct?
11:12:27 A. Yes, it is.
11:12:28 Q. Okay. And as part of that determining step, the delays have to be determined between each of said sound sensors and an origin of said array of sound sensors. Do you see that?
11:12:45 A. That's correct.
11:12:45 Q. So when the device is in operation, it has to determine a delay between each of those seven microphones, in the

example we're looking at, and an origin of the array of microphones; is that correct?
A. That's correct.
THE COURT: Let me -- let me interrupt, counsel. Let's bring the easel out to the corner of the ELMO.
MR. HADDEN: Sure.
THE COURT: That way the jury can see it, opposing counsel can see it, and the witness can see it.
MR. HADDEN: Can we blow up on the screen just the determining steps?
Q. (By Mr. Hadden) And so going back to -- going back to these steps, right?
So we have to perform this determining on the device when it's in operation, correct?
A. Yes. And it is.
Q. And -- and that determining has to be a -- those delays have to be a function of several things, right?
A. Yes.
Q. And they have to be a function of the distance between each of the sound sensors or microphones and the origin --
A. Correct.
Q. -- right?
A. Correct.
Q. And when we say "a function of," that means that those things have to be an input to some determination step that

11:14:28 is performed on the device in order to output these delays;
11:14:28 isn't that right?
11:14:33 A. Well, you say it has to be an input. It has to be a
11:14:36 part of the delay determination.
11:14:39 Q. So in determining -- in this step of determining a
11:14:44 delay as being performed on the device, part of that
11:14:48 determination has to use each of these three items we've
11:14:53 highlighted or pulled out as -- as separate lines, right?
11:14:58 A. Well, you've identified -- you've talked about the
11:15:00 identification of the separate lines for the distance, the
11:15:02 angle, and the azimuth?
11:15:04 Q. Correct.
11:15:05 A. Yeah. All three of those -- and as I've looked at the
11:15:10 code, all three are -- are used as a part of that
11:15:13 determining.
11:15:13 Q. And, again, going back to what these delays are as we
11:15:16 saw in Figure 5, right? In a circular microphone array, if
11:15:21 you have a sound wave coming in from a particular
11:15:25 direction, there are going to be -- that sound wave is
11:15:29 going to hit those microphones at different times, right?
11:15:33 A. Yes, depending upon where it is coming in from.
11:15:38 Q. So this delay is -- or these delays that are being
11:15:41 determined in this step are going to be different for the
11:15:45 different microphones, depending on the direction of that
11:15:50 target sound source, right?

A. Not correct. When you say they're going to be different depending on the direction, for the most part that's true. But if you have -- if you have a sound signal that's coming in tangential at the center of the microphones from the top, it's going to be equally seen by all at the same time, so there's no delay.

Q. That's --

A. The delay is the same.

Q. That's not correct if you have a circular microphone array, because there's going to be one at the top and one at the bottom. That would be true if you have a linear array, right?

A. No. If you -- if you have a system like this where the microphones are at the top and you have a signal tangential coming into the top, all microphones are going to see that at exactly the same time. So your premise that the delay will be different is not necessarily correct.

Q. You're correct. But the language accounts for that, because it talks about the target sound signal being a two-dimensional plane, right?

A. That's correct.

Q. So your example is not in a two-dimensional plane; isn't that correct?

A. That would be a different example, that's correct.

Q. Right. So if we focus on the claim language where we

11:17:00 are in a two-dimensional plane, those delays that are -- are determining are going to vary from microphone-to-microphone?
11:17:06 A. With that qualification, the answer is yes.
11:17:10 Q. Right. And so when we determine these delays, we're going to come up with different numbers for each of the microphones, right?
11:17:19 A. When you look at -- when you consider all three parameters that we're talking about in terms of the distance, the angle calculation, and the azimuth angle, when you look at all three of those portions together for the determining a delay, yes, there will be a different delay that is determined or sensed for each one of the microphones.
11:17:44 Q. Well, it's determined, right? The claim requires determine, right?
11:17:49 A. Correct.
11:17:49 Q. So we're going to do some determining, some process on the Echo device, and it's going to use the angle to the target that we're trying to listen to, and it's going to output some delays that are going to be different for the different microphones, and we also know that those delays have to be represented in terms of samples. Isn't that right?
11:18:13 A. The -- the -- you said the output. There has to be a

11:18:17 determination based upon those three parameters, and the claim does go on to say that where the delay is represented in terms of number of samples, that's correct.

11:18:35 Q. Right. So instead of a time delay, we're going to have some count, some number that corresponds to the number of those 16,000 hertz sample of rate that we're going to have that will correspond to that time delay, right?

11:18:51 A. Right. And as I described on the Amazon products using the Fast Fourier Transform, using the delay techniques that are used there in terms of time but put in a Fourier transform array for frequency, yes, then you end up with counts, you end up with a sample.

11:19:10 Q. Okay. We'll get to that.

11:19:12 MR. HADDEN: Let's look at -- can we look at the slides, Mr. Berk, starting with the Biscuit beam diagram?

11:19:25 Q. (By Mr. Hadden) Now, I think you referenced this slide in your expert report in this case, didn't you, Mr. McAlexander?

11:19:28 A. I recall that to be the case, yes.

11:19:30 Q. Right. So for the Biscuit, which is one of these code names for one of the -- one of the Echo products, that has six beams, right?

11:19:44 A. Correct. Seven -- seven microphones, six beams.

11:19:48 Q. So that is -- that works with some of those Doppler diagrams that you put up during your testimony, right,

11:19:55 which also had seven microphones and six beams, right?
11:19:59 A. That is correct.
11:20:00 Q. And these beams that we see here, those are defined
11:20:03 when the device is made, right?
11:20:08 A. What do you mean by "defined when the device is made"?
11:20:14 Q. Well, those directions that we see there, those are put
11:20:17 into the device when the device is made, right?
11:20:19 A. Well, I'm hoping that the source code defines what is
11:20:22 going -- how the system operates.
11:20:25 Q. Well, so do you agree, Mr. McAlexander, that those six
11:20:33 beams are formed when the device is created?
11:20:36 A. No, the -- the configuration for the beams is dependent
11:20:41 upon, in part, the -- the weighting numbers that are
11:20:46 provided and put into the code when it's actually
11:20:50 implemented in the device. But the beams are actually --
11:20:54 the formation of the beam and the selection of the beam are
11:20:56 done when the spoken word occurs.
11:20:58 Q. Those directions that we see here in the diagram are
11:21:00 defined by the beamformer coefficients; isn't that right,
11:21:04 Mr. McAlexander?
11:21:04 A. The beamformer coefficients are going to be a -- a
11:21:10 significant part of the formation of the beampattern, that
11:21:17 is absolutely correct.
11:21:17 Q. Okay. And those beamformer coefficients are loaded
11:21:20 onto the device in a configuration file when the device is

11:21:27 made and before it is sold; isn't that correct?
11:21:28 A. Yes. These are -- as I mentioned before, the weighting is done by MATLAB or COMSOL, and those inputs are provided as a part of the source code download, and so they are in that device. And then when the determining step is performed, that's when you actually have a target signal of which you are to then evaluate.
11:21:45 Q. We'll get to the target signal.
11:21:47 Those directions that we see here in the figure --
11:21:50 THE COURT: Just a minute, Mr. Hadden. We're getting to the target signal, we'll get to that in a minute, those are sidebar comments.
11:22:00 MR. HADDEN: Apologize, Your Honor.
11:22:01 THE COURT: You need to communicate with the witness in questions.
11:22:05 Q. (By Mr. Hadden) These directions that we see here in the figure, those don't change, do they, Mr. McAlexander?
11:22:10 A. The directions?
11:22:14 Q. Right. Those red circles that we see are oblongs that show the beam directions in the figure on the screen. And those directions do not change after the device is made, do they?
11:22:24 A. The directions themselves do not change in terms of the beamformation. Of course, what is shown here is the primary lobe. Understand there are side lobes to this.

This is just a graphical representation of the primary lobe of the beam.

Q. And those directions and those red oblongs that we see here don't change because the coefficients on the device don't change?

A. That's not correct, because based upon the azimuth of the incoming signal, you will select according to that.

Q. The shape of the beams that we see here in those directions are fixed by the coefficients that are stored on the device, aren't they, Mr. McAlexander?

A. Again, this is a representation of only the primary lobe. There are other lobes here. This is not a full representation of what is there.

Q. The directions that are shown here are fixed by the coefficients on the device, and those coefficients don't change, do they, Mr. McAlexander?

A. The coefficients that are instantiated in the MATLAB or COMSOL code that has been downloaded on to the device, excuse me, that identifies the -- the particular pattern of the beams, and then the pattern of the beams is -- is -- is going to be massaged based upon the incoming azimuth of the signal.

Q. So does this figure change or doesn't it?

A. Well, this figure doesn't change.

Q. And the coefficients that define these directions,

those don't change; they're fixed and stored on the device. Isn't that right?
A. Well, remember, you have -- you have beam coefficients that are -- from the MATLAB that -- that are from simulations that simulate all azimuth angles and all elevation angles.
So you have a conglomerate of a major number of coefficients that are downloaded, but what you then load is based upon the azimuth of the incoming beam at the time it comes in.
So it doesn't -- it doesn't provide this pattern every time for every single beam. The azimuth is going to determine that.
MR. HADDEN: So let's look -- can we go to the next slide? Can you just blow up the top part maybe a little bit?
Q. (By Mr. Hadden) You showed, I think, this representation during your direct testimony. Do you recall that, Mr. McAlexander?
A. Yes, I do.
Q. Okay. And this is for the Doppler products we talked about that have seven microphones and six beams, right?
A. That is correct.
Q. Now -- and as you explained, on the left, coming in, we have the signals from the microphone array; isn't that

right?
A. Correct, that's the seven signals in this particular example coming in at Send-in at -- at the frequency of 16,000 hertz per second.
Q. And then we have -- we do some filtering, we have the microphone calibration block; do you see that, Mr. McAlexander?
A. Yes, sir, I do.
Q. And we have a high pass filter block. Do you see that, Mr. McAlexander?
A. Yes, I do.
Q. And then we get to this fixed beamformer; do you see that block?
A. I do.
Q. Is that where the seven microphone inputs are used to form the six beams?
A. I don't agree with your characterization. The seven microphone inputs are provided as an input for the determination of the beamforming, and that's an input into the -- into the summation -- delay-and-sum algorithm.
Q. And so you agree that we have outputs from seven microphones coming into this fixed beamformer block, and we have six beams coming out, right?
A. That is correct, yes.
Q. And the conversion of those seven microphone inputs

into those six beams is done using those -- what you're calling MATLAB coefficients, those coefficients that are calculated offline and then stored on the device when it's built, right?

A. For the most part, you're correct. Again, the seven -- the six beams represent the -- the beampatterns that are arranged according to the seven microphones. That's already put into the architecture of the device.

The input to the fixed beamformer is the incoming signal. And so there's going to be sampling of that signal, and that provides an input to the beamformer.

So the beamforming is going to take those inputs, it will use certainly coefficients that have been loaded on that device, and selectively determine the beamforming based upon the azimuth direction and the delay calculations that are occurring at that time.

Q. Is it your testimony, Mr. McAlexander, that the determining step, that we have up on that board, is performed in this fixed beamformer block that we see on the diagram?

A. The determining step is determining delay and the origin, and it's for the purpose of enabling beamforming. And so, yes, the delays are calculated in the Fast Fourier Transform area based upon the sample inputs, which I indicated was the M -- FFT. Those sample inputs are

11:28:14 provided in there, transformed into the FFT algorithms, and
11:28:20 then the delay calculations using a filter-and-sum are
11:28:23 performed by the beamforming algorithms.
11:28:26 Q. So it's your testimony that the determining step is
11:28:29 performed in the fixed beamformer block that we see in this
11:28:33 diagram; is that your testimony?
11:28:35 A. It's in part done -- done there, yes. It's in the
11:28:39 formation of the beams, and that's what the determining
11:28:42 step is about is determining delay, and the delay is
11:28:46 represented in numbers of samples. And the last part of it
11:28:49 is: Wherein said determination of said delay enables
11:28:53 beamforming.
11:28:54 Q. I understand --
11:28:56 A. That's what's done.
11:28:57 Q. Is it done in this block, or is it done somewhere else?
11:29:01 A. Well, the block diagram is an overview at a very high
11:29:04 level, and so parts of it are done there, yes.
11:29:08 Q. Are there parts done in another block on this diagram?
11:29:10 A. I have indicated in the code where that is done, and I
11:29:14 believe it is a part of the beamforming algorithm.
11:29:18 Q. So -- and that code, that super directive beamforming
11:29:25 file that you put a few lines of code up with your counsel,
11:29:28 it's your testimony that that code is being performed
11:29:32 within that fixed beamformer block that we see here?
11:29:35 A. The fixed beamformer block defines a -- it's a

11:29:39 functional block. This is where beamforming occurs. When you go look at it, it's -- it's the beamforming algorithms that are on the DSP, the digital signal processor. That is what it's actually doing. It's the execution of that code.
11:29:53 Q. In this functional diagram, the execution of that code is shown as being part of this fixed beamformer block, right?
11:30:00 A. That I agree with.
11:30:02 Q. Okay. And when we're looking at this diagram, again, the output of the fixed beamformer block is six beams, right?
11:30:14 A. That's correct, that's beam -- it's forming the six beams.
11:30:18 Q. Okay. At this point, there has been no selection of a particular beam, right?
11:30:22 A. That's done by the main beam selector block, to the right.
11:30:26 Q. Right. So that happens after the beams are formed in the fixed beamformer block, right?
11:30:33 A. Well, in this particular one, yes. The -- the actual selection is done by the main beam selector code, and that's under the control of basically the fixed beamformer output, as well as the voice activity detector which also is a controlling part of that code.
11:30:52 Q. Sure. But the fixed beamformer outputs six beams. At

11:30:58 that point no beam has been selected; isn't that right?
11:31:00 A. No, that's -- I mean, that is correct, yes, the six beams are provided. Then, based on some noise reduction into the beam selector, and that's where the actual one of six is selected.
11:31:14 Q. Right. So when the fixed beamformer is taking those seven microphone outputs and applying those pre-stored weights to form those six beams, it's not using an angle or a direction to the target sound signal, is it?
11:31:32 A. That's not correct, because part of the -- I had indicated that in the code, there's an end location which is identified -- mLocM, m-L-o-c-M, that is ascribed based upon the geometric arrangement of the microphones. There's also a tune that is done, and that is in the actual part of the beamforming.
11:31:56 Q. But --
11:31:58 A. And -- and then the selection actually goes into the -- the other -- the other part of the algorithm of a function.
11:32:04 Q. Right, but --
11:32:05 A. So --
11:32:05 Q. -- there is no determination of what the target sound source is or where it is until the beam is selected, right?
11:32:13 A. Well, the determining is determining a delay.
11:32:17 Q. I understand.
11:32:17 A. I don't know what you mean by determining a sound

source. It's not in the claim.
Q. The claim requires that as part of the determining step, you have to use as part of our function an azimuth angle between said reference axis and said target sound signal. Do you see that, Mr. McAlexander?
A. Correct. And that is done by the Amazon products.
Q. Okay. Now, focusing on the fixed beamformer, which you say is where this determining step is performed, when we're in the fixed beamformer, we don't know where the target is or what it is, right? We're just -- we just have the microphone outputs and our stored weights. We haven't selected a beam.
A. It doesn't say selection. It says determining, sir, and that's what's done.
Q. I understand. But the determining requires knowing the angle of the target sound source?
A. And I don't know what you mean by knowing. Are you applying a human characteristic to the product? It's determining. And determining is done by execution of code, considering the different parameters that are required, which is the angle, the distance, and the azimuth.
Q. Right. So if we're going to have -- determine something as a function of an azimuth angle to the target, don't we have to have that azimuth angle to the target as an available input to our function, in order to determine

11:33:41 something using it?
11:33:41 A. It is, because that azimuth angle is a part of the determination because you have the delay inputs. That's how you perform the determining so that you can eventually select from that determining step the -- the actual beam that you're designing.
11:33:59 Q. So if I'm going to be doing my determining step that you say it happens in this fixed beamformer block, the fixed beamformer block has to at least have access to the azimuth angle between the reference axis and the target sound signal, right?
11:34:19 A. It has to have access to it, and I understand that the -- that the adaptive beamformer includes this fixed beamformer and the fixed beamformer does the delay calculation and the azimuth part of it. Then does the -- the steering.
11:34:34 Q. So, just to be clear, if I'm going to determine my delays in the fixed beamformer block, as you say happens, the code that is operating in that fixed beamformer block has to have access to this azimuth angle to the target sound source, right? It's got to be able to use that angle; isn't that right?
11:35:01 A. You keep going back to this block. The block is a high diagram -- a high-level diagram. What I've indicated is some of the functions are built in there, but the execution

11:35:07 of the code is where it is located.
11:35:13 I've identified where the super directive
11:35:15 beamforming is occurring in the code and identified that
11:35:17 the -- the considerations for that Fast Fourier Transform
11:35:25 algorithm for the -- for the filter-and-sum is, including
11:35:29 those parameters, including azimuth.
11:35:30 Q. Well, let's look at that actually.
11:35:33 MR. HADDEN: Can we get Defendants' 318, please,
11:35:36 Mr. Berk -- 319, I'm sorry.
11:35:47 Q. (By Mr. Hadden) Do you see the Amazon document
11:35:50 entitled Review of Doppler Beamformer and Acoustic Echo
11:35:54 Canceler? Do you see that, Mr. McAlexander?
11:35:56 A. I see it. Let me pull it up here. Okay. I see the
11:36:09 document.
11:36:10 MR. HADDEN: And if we look at Page 6, please,
11:36:13 Mr. Berk. Can we blow up --
11:36:18 Q. (By Mr. Hadden) There's a Heading 5, Fixed Beamformer.
11:36:24 Do you see that?
11:36:25 A. Yes, I see that.
11:36:32 Q. And if you look at the last sentence in that
11:36:34 paragraph -- and, again, this is an Amazon document that
11:36:38 describes that fixed beamformer block in Doppler that we
11:36:42 were looking at just a minute ago, isn't it,
11:36:45 Mr. McAlexander?
11:36:46 A. Okay.

11:36:46 Q. And if you look at the last sentence in this paragraph, it says: The term "fixed" in the fixed beamformer implies that for each look-direction, the filtering process is fixed and does not vary with the time or input signal.
11:36:51
11:37:00
11:37:04
11:37:08 Do you see that?
11:37:08 A. That's correct. That's part of the definition of a fixed beamformer.
11:37:14
11:37:14 Q. Okay. And you don't dispute that that accurately describes the way the beams are formed in these Echo products, do you?
11:37:16
11:37:17
11:37:18 A. No, I don't dispute that at all.
11:37:20 Q. And then the next paragraph goes in -- goes on and it says: The FBF algorithm is implement -- implemented using a filter-and-sum structure wherein we apply frequency-domain beamformer weights to the FFT of each microphone signal.
11:37:22
11:37:29
11:37:32
11:37:38
11:37:39 Do you see that?
11:37:39 A. Yes, I do.
11:37:41 Q. Okay. And the filter-and-sum structure is what you talked about with your counsel, right, where the coefficients are applied to the different microphone outputs and they are summed up and that's how you form the beam. Right?
11:37:51
11:37:54
11:37:57
11:38:01
11:38:03 A. That is correct.
11:38:03 Q. Okay. And this explains that those weights and -- are

11:38:09 applied in the frequency domain. Do you see that?
11:38:13 A. I see that.
11:38:14 Q. And as you explained, in a frequency domain, instead of having numbers that correspond to samples of the wave at different times, you have numbers that correspond to the frequency components of the wave; isn't that right?
11:38:36 A. That is correct.
11:38:37 Q. Okay. And so Amazon on the Echo does this filter-and-sum in this frequency domain. You understand that, right?
11:38:51 A. Yes, that's correct.
11:38:55 MR. HADDEN: Can we scroll down to the figure, please, Mr. Berk?
11:39:01 Q. (By Mr. Hadden) And this is a diagram that shows that filter-and-sum process, that the device performs in the frequency domain. Do you recognize that?
11:39:09 A. Yes, I do.
11:39:10 Q. Okay. And is it in this diagram where you're saying the delays are calculated or determined?
11:39:15 A. That's part of the delays are calculated -- are determined here, yes.
11:39:20 Q. Okay. So let's -- just so we have all the description of this --
11:39:26 MR. HADDEN: Can we go to Page 12 of this document, Mr. Berk, and blow up the FBF cycle analysis

11:39:40 chunk there? Okay. That's fine.
11:39:57 Q. (By Mr. Hadden) So if we look at the corresponding description in the document, Mr. McAlexander, that says FBF, that's fixed beamformer, right?
11:40:09 A. That is correct.
11:40:10 Q. It says FBF operates on 256 frequency bins, and it has 512-point FFT. Do you see that?
11:40:22 A. Yes, I do.
11:40:24 Q. So that means that for each frame of the audio data that we are processing through this filter-and-sum algorithm, there are 256 of these FFT coefficients, right?
11:40:44 A. That is what it says, yes.
11:40:46 Q. Okay. And -- and we have those 256 coefficients for each microphone output, right?
11:41:01 A. That is also correct.
11:41:03 Q. And then it goes on and says: Hence FBF beam weighted factors are computed offline and stored as a three-dimensional complex array.
11:41:16 Do you see that?
11:41:17 A. Yes, I do.
11:41:18 Q. Now, those FBF beam weighted factors that are computed offline, those are those coefficients you talked about that were computed using that MATLAB code, right?
11:41:29 A. MATLAB and/or COMSOL, yes.
11:41:32 Q. Right. And just to be clear, MATLAB or COMSOL, those

11:41:36 are programs that engineers and scientists at Amazon use to perform mathematical simulations; isn't that right?
11:41:46 A. That's correct. And it's not an Amazon product. I mean, MATLAB is a general purpose --
11:41:52 Q. Sure.
11:41:53 A. -- provider and so they use that to perform calculations and make a determination on simulations of exactly where the -- all the possible combinations are so that they can account for it on the device.
11:42:06 Q. Sure. But that MATLAB code or that COMSOL code, that is not code that is running on any of these Echo products, right?
11:42:14 A. The MATLAB code is not running on the Amazon product, just like none of the source code is running on the Amazon product. It's what is compiled and provided to the device. That's what's executed.
11:42:26 Q. Well, but there's nothing from MATLAB that is executed on the Echo device, right?
11:42:33 A. Disagree. They determine coefficients.
11:42:36 Q. Well, the coefficients are just numbers? Those are not programs that are executed on it?
11:42:43 A. Well, I don't want to confuse the jury. The -- the MATLAB program is not running on the -- on the device. No, that's done separately to determine the -- the environment of coefficients that can be used. It's the coefficients

11:42:55 that are downloaded.
11:42:56 Q. Okay. Thank you.
11:42:57 So those FBF weight factors here that are computed offline, those are described here as being a three-dimensional complex array.
11:43:09 Do you see that?
11:43:09 A. I do, yes.
11:43:13 Q. Okay. And described as a complex array because each one of those weights is actually what's called a complex number; isn't that right?
11:43:24 A. That is correct. Both real and imaginary.
11:43:28 Q. Right. So it has a real part, and it has an imaginary part, right?
11:43:32 A. That is correct.
11:43:33 Q. And it's not called imaginary because it doesn't exist; it's called imaginary because it is multiplied by the square root of minus 1, right?
11:43:43 A. I'm glad you clarified that.
11:43:46 Q. Am I correct?
11:43:47 A. Yes.
11:43:48 Q. Thank you.
11:43:48 So those are just numbers, right? They're numbers that have this real piece, and they have this imaginary piece, right?
11:43:56 A. Correct.

11:43:56 Q. And those numbers are not delays, right?
11:43:59 A. They are numbers that are coefficients determined as -- as the results of these calculations.
11:44:06 Q. But -- but those -- just starting with the numbers that we're talking about, the -- this complex three-dimensional array. The numbers in that three -- complex three-dimensional array that's stored on the device, those are not delays?
11:44:24 A. They are numbers that are used as coefficients in the calculation.
11:44:28 Q. Right. So those are just numbers. And they're also not a number of samples, are they?
11:44:35 A. The number of samples is based upon the input target signals.
11:44:43 Q. But just to be clear, those complex weights that are stored on the device are not a number of samples, right?
11:44:49 A. No, that's -- that is correct. The coefficients are a part of a -- they are used in part of a formula for determination of a delay.
11:44:59 Q. Right. So then we take those numbers, those -- that complex three-dimensional array. And if we look at this diagram -- so let's just start with --
11:45:16 MR. HADDEN: Maybe if you could blow up the diagram on the left just a little bit more, Mr. Berk, so we can see better. Okay. Great. Perfect. Perfect.

11:45:27 Q. (By Mr. Hadden) Okay. So if we can -- if we look at what's going on here. So we have this 7 channel coming in from the left; do you see that, Mr. McAlexander?
11:45:43 A. I do, yes.
11:45:44 Q. And that is the outputs from the microphones after they've gone through this high pass filter, correct?
11:45:51 A. That would be correct.
11:45:52 Q. And each -- what comes in there is just a little chunk of the microphone output for a very short time period, eight milliseconds; isn't that right?
11:46:08 A. I'm not sure what you mean by coming in from the microphone. The microphone is going to be sending out as signals are coming in. So I don't know about the time frame that you're putting this.
11:46:20 Q. Well --
11:46:21 A. I don't understand your question.
11:46:21 Q. Sure. When we were performing the calculation, this filter-and-sum calculation that you talked about and it's shown in this diagram, that is done for one frame of audio at a time; isn't that right?
11:46:38 A. That is correct. I thought that -- what I understood what you said you were talking about coming in from the microphone which would not be time limited, but the Fast Fourier is framed, yes.
11:46:50 Q. Right. So the little arrow that comes into the FFT

block, that comes in as a frame at a time, and each of those frames is represented by 128 samples, right?
A. That's correct. The incoming signal is stored as -- in a buffer, and then they are framed into the Fast Fourier Transform, yes.
THE COURT: Wait a minute. Mr. McAlexander. "That's correct" was a complete answer.
THE WITNESS: Okay.
THE COURT: And then "the incoming signal is stored in a buffer," that was not called for by the question. Try to limit your answers to the questions asked.
THE WITNESS: All right, sir.
THE COURT: Let's continue, counsel.
MR. HADDEN: Thank you, Your Honor.
Q. (By Mr. Hadden) And then for one of those 128 sample frames that go into the FFT block, what we get out is one of these 256 complex numbers corresponding to those 256 different possible frequencies, right?
A. I agree with that.
Q. Okay. And -- and we get one of those for every one of these lines X1, X2, down to X7, right?
A. I believe that is correct, yes.
Q. So those X1, X2 to X7, those represent the Fast Fourier Transform for each of those frames for each of the seven

11:48:16 different microphones, right?
11:48:18 A. Correct.
11:48:18 Q. And then -- and we do that -- we repeat that process where we use the Fast Fourier Transform numbers six times for each of the six beams we're trying to form, right?
11:48:38 A. Correct.
11:48:39 Q. So this -- the first chunk of this is just dealing with one beam, the first part of that diagram, and then we have -- we kind of skip the intermediate beams and we go down to the last Beam 6 and repeat it here in this diagram, right?
11:48:56 A. That's what the diagram says, yes.
11:48:58 Q. Right. So let's just focus on Beam 1 for now.
11:49:02 So we have these 256 complex numbers corresponding to each of these lines, X1, X2, X7. And if we start with X1, we have those 256 complex numbers, and then we have this circle with the X through it. Do you see that?
11:49:22 A. Yes.
11:49:23 Q. And that's where we multiply the Fast Fourier Transform numbers by these pre-stored beam weighting coefficients that are stored on the device, right?
11:49:38 A. Correct.
11:49:38 Q. And there'll be one of those numbers for each beam and each microphone, right?
11:49:48 A. Coefficients would be for each microphone, yes.

11:49:54 Q. Right. So this says W1_1(f). So what that means is that would be the pre-stored coefficient for Microphone 1, for Beam 1, and for Frequency f where f goes from 1 to 256, right?
11:50:14 A. Correct.
11:50:14 Q. So that little circle there is really showing 256 different complex multiplications, right?
11:50:26 A. I believe that is accurate, yes.
11:50:28 Q. Okay. So is there a delay being calculated in that multiplication?
11:50:41 A. It's based upon the input signal itself, so there's delay that's being -- that's being determined in this particular area, yes.
11:50:44 Q. So it's your testimony that when I multiply my Fast Fourier output for Microphone 1 by the -- by complex weight, that that is determining a delay?
11:51:07 A. That is part of the determining a delay, yes, because it is based on the input signal, and that input signal is then evaluated based on the samples in the Fast Fourier Transform domain.
11:51:22 Q. So, just to be clear, your testimony is X1, that top line where we have the Fourier Transform outputs, if we multiply those by the fixed weights, that circle, at that point, I have determined the delay that's required by this Claim 1?

11:51:40 A. The -- the full determination of that is done -- the answer is, no, that's not complete because azimuth is not a part of this.

11:52:04 Q. Right. So at this point, just to be clear, in this whole process that we show here in this filter-and-sum process, nowhere are we using an azimuth angle to a target sound source, are we?

11:52:19 A. In this particular portion, that is correct.

11:52:21 Q. So is there -- if we look at the rest of this diagram, we are computing these filter and weight sums for each beam, right?

11:52:48 A. Yes, that is correct, for each beam.

11:52:50 Q. And they're all done the same way except for those pre-stored weights, the Ws with the 1, 2, or 6, 1, because those are different because those are specific to a particular beam direction?

11:53:09 A. That is also correct at the time the beam is coming in.

11:53:14 Q. So if -- the output of this is going to be a collection of complex numbers, one for each beam, right?

11:53:30 A. That would be correct, yes.

11:53:31 Q. Okay. And is that the complex number for, say, Beam 1, is that a delay?

11:53:40 A. It's a different number than the other beam. As I said, the -- you're asking, is that a delay? It is a delay -- it's a part of the delay, but the azimuth portion

is not part of this as yet, based on this diagram.
Q. How can it be a delay if the delay has to be a function of the azimuth angle and we don't have the azimuth angle anywhere in this diagram?
A. I just said, the -- it's not a completed delay because you have added several parameters here, but you haven't added in the azimuth in this diagram.
Q. Well, what part of that complex number is a delay?
A. This is the portion of the delay that will eventually create the delay based upon the azimuth addition to it. This -- this is a part of it. It's in the process. It's not complete.
Q. So is that complex number that is output from this process a number of samples?
A. Yes. The sample is input, it's in the Fast Fourier realm, yes, it's sample output.
Q. No. The question was whether it is a number of samples. The claim requires the delay is represented as a number of samples?
A. Yes, it's -- excuse me. It's a number of samples, yes.
Q. Well, so how can a complex number be a number of samples?
A. You got different numbers.
Q. The complex number -- a number of samples has to be an integer; you agree with that, don't you?

A. A number of samples can be used as an integer, and if you look at the formulation, you'll notice that the imaginary component is discarded.
Q. Even if the imaginary component is discarded, this number is not an integer?
A. It's outputted as a number of samples for conversion back into the -- into the time domain.
Q. Well, a Fourier Transform coefficient is, by definition, not in the time domain, you understand that, right?
A. That's correct. I said it's output to the time domain.
Q. Right. So all of these Fourier Transform coefficients are calculated using initially the same number of samples, right? 128?
A. In -- per frame, yes.
Q. Right. And all of these beams use the same 128 bit frame, right?
A. That is correct, per -- per frame.
Q. So are you saying that the 128 bits is the number of samples in this delay?
A. No, because it loops. It's taking sample slices, and it will continue and then -- sum of the slices. That's what the summation does. Each one is 128, but then it will continue to sum until it finishes evaluating the incoming beam.

Q. And all of these beams, then, are receiving the same number of samples from each of the microphones, right?
A. Yes, that's correct.
Q. So how could the delay, if it's the number of samples that are coming in from the microphone, vary from beam-to-beam or microphone-to-microphone?
A. I'm not sure what you mean how can they vary.
Q. Well, we already discussed, right? The whole point of this determining a delay calculation is you have to determine a delay between each of the microphones and an origin, and that delay has to be represented as the number of samples?
A. That's correct, represented.
Q. Right. So if I have a delay, instead of saying it's .5 seconds, I would say it's -- it's some number of those 128 frames or samples, right?
A. It would be some number of a number of a frame -- a number of frames that have been evaluated.
Q. But nothing in this Fourier Transform process where we're taking samples and converting them to frequencies depends at all on the distance between the microphone and the origin or the angle between the microphone and the origin or this azimuth angle because we don't even know what the azimuth angle is at this point, right?
A. Well, you've got two -- two prongs to that question.

Yes, the microphone array location is already identified and known, so the weighting factors are ascribed based upon that arrangement, so the location is known.

Q. But the question, sir, was, with respect to the number of samples. You're saying the number of samples is somehow related to this Fast Fourier Transform, but that Fast Fourier Transform has nothing to do with the weighting factors. We apply the weighting factors later when we get to the Ws. Right?

THE COURT: Is that a question, counsel? I heard about five statements, and then a question at the end.

MR. HADDEN: I'll rephrase it. I apologize, Your Honor.

Q. (By Mr. Hadden) You talked about the geometry somehow being in the weighting factors.

Isn't it true, sir, that when we're doing the Fourier Transform and we're creating these Xs, we haven't dealt with anything regarding the geometry of those microphones? We're just dealing with the outputs and these 128 samples per frame; isn't that right?

A. No, that's not correct.

Q. So are you saying that the number of samples from each microphone varies depending on its angle with respect to the reference axis?

A. There's nothing in the claim that says the number of

samples varies in accordance with; it's just represented by a number of samples.
Q. Okay. But if the delays are going to be different, then the number of samples has to be different, right?
A. No.
Q. No?
A. It doesn't have to be. It can be.
Q. If you're going to have two delays that are not the same and each of them is represented by or in terms of a number of samples, don't the number of samples have to be different between those two?
A. The eventual result would result in a factor that when it's converted back to a time domain would be representative of a difference, yes.
Q. But nowhere in this diagram or this process, have we determined anything in terms of a number of samples other than a fixed 128-sample frame, right?
A. Again, that's per frame, and then you -- you -- the different -- the value is different based upon the weighting factors. So it's going to be a difference.
THE COURT: Let me interrupt here, gentlemen.
MR. HADDEN: Sure.
THE COURT: It's clear this cross-examination has some additional time to go, and we're at the noon hour now. So we're going to recess for lunch.

12:00:53 Ladies and gentlemen of the jury, if you'll close your notebooks and take them with you to the jury room, I'm told by the clerk that your lunch is there waiting for you.

12:01:04 It's 12:01 by the clock here on the bench. We'll try to reconvene as close to 12:45 as possible. Follow all the instructions I've given you about your conduct, including, of course, as you would expect me to remind you, not to discuss the case among yourselves. And we'll be back at that time or close thereto, to continue with the Defendants' cross-examination of this witness.

12:01:28 The jury is excused for lunch at this time.

12:01:32 COURT SECURITY OFFICER: All rise.

12:01:32 (Jury out.)

12:01:32 THE COURT: The Court stands in recess for lunch.

12:01:56 (Recess.)

CERTIFICATION

I HEREBY CERTIFY that the foregoing is a true and correct transcript from the stenographic notes of the proceedings in the above-entitled matter to the best of my ability.

/S/ Shelly Holmes ____________ 10/5/2020
SHELLY HOLMES, CSR, TCRR Date
OFFICIAL REPORTER
State of Texas No.: 7804
Expiration Date: 12/31/2020